Can an inhibitor of a multidrug pump become a substrate?

Ella Hutchison

Thesis submitted to the University of Nottingham for the degree of Master of Research

COVID-19 impact statement

The research for this thesis was performed October 2019-March 2020 followed by a 7-month interruption to studies and then finished November 2020-May-2021. This was due to closure of the university labs and the break also had the impact of additional work for lab shutting down and lab set up (freezing cell lines, thawing and reviving cells etc, all adding time and further disruption).

After returning in November, safety measures limited lab occupancy to 5 people and only permitted coming into the lab when absolutely necessary for lab work. This was challenging because working from home for all non-laboratory activities required me to pre-plan all lab work from home which was difficult without access to all the lab resources e.g., checking stock of reagents. This also resulted in reduced flexibility for working in general. Additional cleaning requirements and social distancing measures both made day to day work more challenging and time consuming.

As an early stage career researcher social distancing was particularly impactful on my learning and productivity as training and supervision were both limited. This required me to work more independently and change my plans, for example I was planning to use confocal microscopy but the machine I was trained on was no longer in use due to COVID-19 restrictions.

Abstract

ABCG2 is an ATP binding cassette (ABC) transporter that is involved in multidrug resistance, particularly anti-cancer drugs such as methotrexate and mitoxantrone. Distinguishing between substrates and inhibitors is important for rational design of drugs that won't be transported by ABCG2 and inhibitors that prevent transport of existing ABCG2 substrates. In this project, the hypothesis being proposed is that ligands with a higher affinity for ABCG2 act as inhibitors and the transient conformational changes required for transport do not occur. Whereas transported substrates have a lower affinity for ABCG2 and the conformational changes can occur. Therefore, if the affinity for a widely used inhibitor, Ko143, was reduced, would it become a substrate that is transported by ABCG2? A series of mutants (T435A, N436A, F439A, S440W, M549E, A397S/V401A/L539A and L405A/I543A/V546A) were designed using cryo-EM structures of ABCG2 bound to MZ29 (a Ko143 analogue) and mutational studies of substrate transport. The aim was to reduce affinity of ABCG2 for Ko143. All mutant proteins, except M549E, were successfully expressed in HEK293T cells and trafficked to the cell membrane. Using flow cytometry, cellular accumulation of fluorescent Ko143 derivatives (Ko143-Cy5 and Ko143-X-BY630) was measured which indicates whether they are exported by ABCG2 or not. Ko143-Cy5 fluorescence was significantly higher in WT-ABCG2 expressing cells compared with the untransfected control and the other ABCG2 mutants, except for A397S/V401A/L539A. Ko143-X-BY630 showed a similar pattern but without significance. The reduction in cellular accumulation of the fluorescent Ko143 derivatives in the mutants could be caused by the reduced affinity leading to ABCG2-mediated transport or increased diffusion out of the cell. Transport of Ko143-Cy5 or Ko143-X-BY630 cannot be ruled out but it is not detectable in this

experiment. Consideration of the data in this thesis alongside emerging structural and functional data from other laboratories will continue to shed light on the interaction of substrates and inhibitors with ABCG2.

Acknowledgements

I would like to thank Dr Ian Kerr for his guidance and support throughout my research project. This support and advice has been especially appreciated in writing this thesis. I would also like to add gratitude for encouraging my development as a researcher during challenging times and having faith in me to contribute to a review article.

I also want to thank Deb Briggs for all her help and patience when teaching me new techniques and giving me advice when running my experiments.

I am thankful to my fellow lab members Dr James Mitchell-White, Joe Morris and Hannan Azmir for their moral support and for all of the advice and help in the lab. I also appreciated and enjoyed our lunch time chats. I would also like to thank Dr James Mitchell-White for his contribution to the monolayer transport assay.

Thanks to Dr David Onion and Nicola Croxall (University of Nottingham Flow Cytometry Facility) for giving me the training, advice and use of equipment for this project.

Also, a huge thank you to my boyfriend, Joe Davies, for being a constant source of motivation and for encouraging me to do a Master's in the first place.

Declaration

This thesis, “Can an inhibitor of a multidrug pump become a substrate?”, is the result of my own work undertaken during my period of registration at the University of Nottingham under the supervision of Dr Ian Kerr. Technical assistance, and collaborations where relevant, has been acknowledged.

Ella Hutchison

Student ID: 20214453

May 2021

Abbreviations

ABC	ATP-binding cassette
ABCG2	ATP-binding cassette subfamily G member 2
ABCG5/G8	ABCG5 heterodimerised with ABCG8
ADP	Adenosine diphosphate
AMP-PNP	Adenylyl-imidodiphosphate
ANOVA	Analysis of variance
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine serum albumin
CFTR	Cystic fibrosis transmembrane conductance regulator
CMV	Cytomegalovirus
Cryo-EM	Cryogenic electron microscopy
DHEAS	Dehydroepiandrosterone sulfate
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
dNTP	Deoxynucleotides
E_1S	Estrone 3-sulfate
E3040S	E3040 sulfate
EC_{50}	Concentration at which a drug evokes half maximal response
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
FSC	Forward scatter
FTC	Fumitremorgin C
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
HBSS	Hank's balanced salt solution
HDL	High density lipoprotein
HEK293S	Human embryonic kidney 293S cell line

HEK293T	Human embryonic kidney 293S cell line
IC_{50}	Half-maximal inhibitory concentration
IgG	Immunoglobulin G
kb	Kilo bases (1000 base pairs)
kDa	Kilo Daltons
LB	Luria-Bertani
NBD	Nucleotide binding domain
NCBI	National Center for Biotechnology Information
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with 0.1% (v/v) Tween-20
PCR	Polymerase chain reaction
PEI	Polyethyleneimine
P_i	Inorganic phosphate
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SSC	Side scatter
SV40	Simian vacuolating virus 40
TBE	Tris/borate/EDTA
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TLC-S	Taurolithocholate sulfate
TM	Transmembrane helix
TMD	Transmembrane domain
UV	Ultraviolet
WT	Wild type
4-MUS	4-methylumbelliferone sulfate
ΔMFI	Change in median fluorescence intensity

Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Membrane transporters

All forms of life have a lipid membrane of some sort to contain their molecular components to allow biochemical reactions to take place. Without membranes, the compartmentalisation necessary for the vast majority of cellular processes cannot occur. Archaea, bacteria and eukaryotes have lipid bilayers which act as a barrier so that what enters and exits the cell can be controlled. In eukaryotic organisms, different cellular processes are facilitated and regulated by their spatial separation in organelles.

Integral membrane proteins have sections that are inserted into the cell membrane and can perform a variety of roles (Figure 1.1). For example, flippases and floppases can alter the shape and lipid composition of the cell membrane by transferring lipids from the inner layer to the outer layer or vice versa. Receptors such as G-protein coupled receptors (GPCRs) can aid in the communication between cells, for example by triggering signalling cascades when a signal molecule binds. Transporter and channel proteins control what enters and exits the cell or organelle. This can include transport of charged ions, for example cystic fibrosis transmembrane conductance regulator (CFTR) exports Cl^- from endothelial cells to control mucus consistency and in neurons Na^+ influx can trigger an electrical impulse. Larger molecules such as peptides, amino acids, nucleotides and phospholipids can also be transported. Transporters can also have a protective effect by exporting potentially harmful xenobiotics back into the lumen of the gastrointestinal tract or away from the foetus when expressed in the placenta (Horsey et al., 2016). This is the role of the membrane transporter, ABCG2, studied in this thesis.

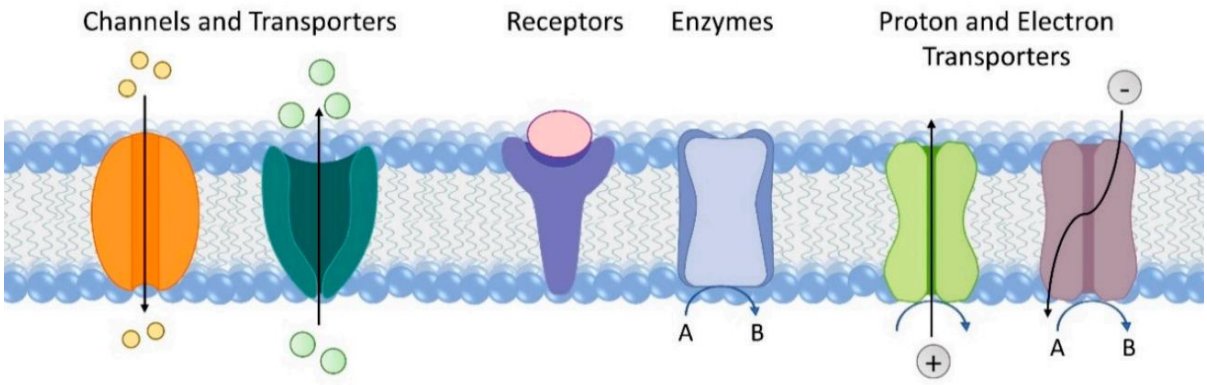


Figure 1.1 Schematic of the cell membrane. The lipid bilayer (blue spheres and tails) contain different types of membrane proteins involved in signalling, enzymology and substrate flux. Image retrieved from (Zhang et al., 2020)

1.1.1 Facilitated diffusion vs active transport

Some molecules can cross the cell membrane unaided because they are small and hydrophobic so can easily diffuse through the hydrophobic lipid bilayer. Movement of these molecules occurs down the concentration gradient, from a more concentrated cytosol to the less concentrated extracellular space, or vice versa. More polar compounds (e.g. water) or ions (e.g. Na^+ or H^+) cannot diffuse across the cell membrane without the help of channels or carrier proteins by facilitated diffusion. These proteins selectively facilitate the diffusion of chemical entities through a more suitable chemical environment. Flux can be controlled, for example voltage gated Na^+ channels undergo conformational changes to open the channel when membrane depolarisation occurs (de Lera Ruiz and Kraus, 2015).

Active transport can be against a concentration gradient and the transported compound is “pumped” in or out of the cell. An important step in the transport cycle of primary active transporters is ATP hydrolysis, which provides the energy for movement against the concentration gradient. An example of proteins that perform active transport are ATP-binding cassette (ABC) transporters which are described in

more detail below. The subject of this project, ABCG2, is part of this family of proteins. The true substrate of active transport proteins is ATP because its hydrolysis to produce ADP and P_i is catalysed by the transporter protein, however, in this thesis, the substrate will refer to the transported compound as opposed to an inhibitor of the pump.

1.1.2 ATP-binding cassette transporters

The ABC family of proteins are mostly membrane proteins: in humans ABCA, ABCB, ABCC, ABCD and ABCG subfamilies are membrane associated but ABCE and ABCF are not (Kerr et al., 2011). The basic structure of an ABC transporter is two nucleotide binding domains (NBDs) and two transmembrane domains (TMDs), however some ABC transporters, named “half-transporters”, must dimerize in order to achieve this formation. The NBDs are highly conserved, even between eukaryotic and prokaryotic transporters of otherwise unrelated function. They contain several motifs that are essential for ATP-hydrolysis, which is what drives the active transport. Figure 1.2 shows that the P loop (or Walker-A motif) from one NBD and the LSGGQ motif (or C-loop) from the other NBD contact the nucleotide, resulting in two ATP molecules bound at two separate points along the NBD:NBD interface (Jones and George, 1999, Hollenstein et al., 2007). The Walker-B motif contains a catalytic glutamate which is capable of performing a nucleophilic attack on ATP via a water molecule (Hollenstein et al., 2007). The TMDs, which span the cell membrane, are less conserved and relate more to the specificity for the transport substrate and perform conformational changes required for substrate transport. TMDs contain a coupling helix (black helices, Figure 1.2) that contact the NBD (Q loops) to couple ATP hydrolysis to the conformational changes required for substrate transport

(Hollenstein et al., 2007). ATP binding and hydrolysis is coupled to inward and outward facing conformations of the transporter TMDs which in the case of exporters, allows binding of substrate on the intracellular side and release on the extracellular side (Hollenstein et al., 2007, Manolaridis et al., 2018).

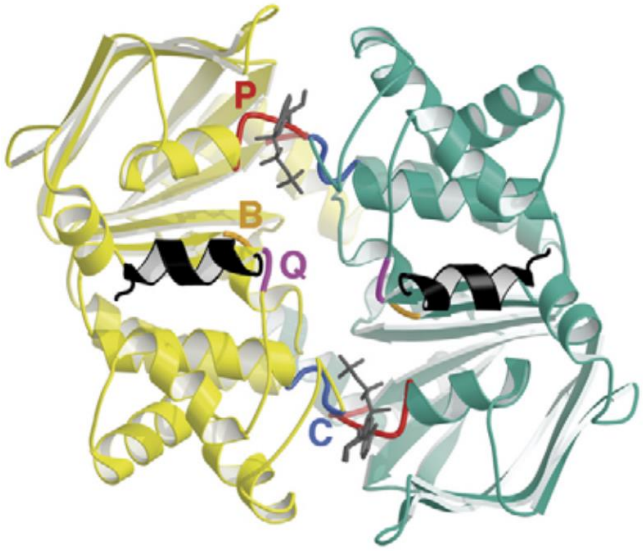


Figure 1.2 The NBDs of the ABC transporter Sav1866. The NBDs of Sav1866 crystallised with bound AMP-PNP, a non-hydrolysable ATP analogue, shown in stick representation. There are 2 bound AMP-PNP molecules per NBD dimer. The NBDs are shown as if looking down from the membrane. The 2 NBDs are in green and yellow and mechanistically important sequence motifs are shown with single letter: red P (P-loop), yellow B (Walker-B motif), purple Q (Q-loop), blue C (LSGGQ or C-loop. Bound AMP-PNP is shown as grey sticks and coupling helices from the TMDs are the short black helices. Figure adapted from Hollenstein et al. (2007)

1.2 ABCG subfamily

1.2.1 Overview

The ABCG subfamily is part of the ABC family of proteins and the 5 members of this subfamily all share a common ancestral gene. Besides the conserved ABC motifs, there is very little protein sequence homology between the ABCG proteins, except G1 and G4 which have 72% homology (Kerr et al., 2011). This plays down to the

TMDs being largely non-conserved, for substrate selectivity (Hollenstein et al., 2007). The ABCG family all transport lipids and, with the exception of ABCG2, have a narrower functionality than other ABC subfamilies (Kerr et al., 2011, Kerr et al., 2021). All ABCG proteins are half-transporters because they only consist of one TMD and one NBD which requires them to at least dimerize to function. This is so that the ATP binding site can form at the interface between two NBDs (section 1.1.2). ABCG5 and ABCG8 form an obligate heterodimer to complete this binding site whereas ABCG2 homodimerises but higher order oligomeric structures have been observed (Wong et al., 2016, Kerr et al., 2011). ABCG1 and ABCG4 can homodimerise but have also been shown to heterodimerise with each other *in vitro* (Hegyi and Homolya, 2016).

ABCG1 has widespread expression and is located in the brain alongside ABCG4 and both contribute to cholesterol and desmosterol efflux to high-density lipoprotein (HDL) from astrocytes (Wang et al., 2008). ABCG5/G8 is involved in limiting absorption of dietary sterols by localising in the apical membranes of hepatocyte canaliculi and gall bladder epithelial cells. ABCG2 has a broad specificity, having a wide range of substrates structurally unrelated to each other. This is useful in its protective role against xenobiotics (Kerr et al., 2011). Its function, structure and mechanism are described in more detail below.

1.2.2 ABCG2 function and clinical implications

ABCG2 is an important protein to study because of its role in multidrug resistance. Its natural role as a defence against environmental toxins means that some drugs fall victim to this multidrug pump. Expression of ABCG2 on apical membranes of epithelial cells of the gastrointestinal tract, liver canalicular membranes and the apical

membrane of proximal tubular cells in the kidneys, contributes to limiting absorption and promoting excretion of xenobiotics (Horsey et al., 2016). Expression in placental syncytiotrophoblasts and in mammary glands have contradictory roles, with ABCG2 pumping xenobiotics away from the foetus and into milk respectively. Along with reduced uptake and increased excretion, ABCG2 contributes to the failure of some anti-cancer drugs by its overexpression in cancer cells. ABCG2 expression has been linked to poor outcome in acute myeloid leukaemia, diffuse large B-cell lymphoma, and lung and oesophageal cancer (Horsey et al., 2016). Overexpression of ABCG2 has also been found in other cancer cells exhibiting a multidrug resistance phenotype, including topotecan-selected ovarian tumour cell line T8 and gefitinib-resistant non-small cell lung cancer (NSCLC) cells (Mo and Zhang, 2012). This is why it is so important to study ABCG2 further, to be able to design drugs that won't be exported by this multidrug transporter. Drugs that have been shown to be transported by ABCG2 include: methotrexate, mitoxantrone, pheophorbide A, topotecans, flavopiridol, imatinib, gefitinib, nilotinib and others (Homolya et al., 2011, Volk and Schneider, 2003, Kapoor et al., 2018, Robey et al., 2005). Multidrug resistance is caused because ABCG2 limits absorption, increases excretion and lowers cellular concentration of these drugs. An endogenous substrate is urate and successful transport by ABCG2 leads to urate excretion. A naturally occurring single nucleotide polymorphism, rs2231142, which results in Q141K mutant ABCG2, is a loss of function mutation causing an increased level of serum urate which is linked to hyperuricemia and gout (Woodward et al., 2009).

1.2.3 ABCG2 structure and mechanism

ABCG2 contains one transmembrane domain (TMD), which consists of six transmembrane helices (TM1-TM6), and one nucleotide binding domain (NBD), which is cytoplasmic (Kapoor et al., 2018). Since two NBDs are required for ATP hydrolysis to occur, dimerisation is required for the protein to be functional. Higher forms of oligomerisation, such as tetramers, have been observed, although the physiological relevance is unclear (Wong et al., 2016). In the TM5-TM6 extracellular loop region (Figure 1.3) there is an intramolecular disulfide bond (C592 and C608), an intermolecular disulfide bond (C603 from each monomer) and a glycosylation site (N596) (Diop and Hrycyna, 2005, Henriksen et al., 2005, Kapoor et al., 2018). The mutation of C603 does not prevent surface expression of the protein or affect its function so presumably the C603 inter-molecular disulfide bond just has a small stabilising effect. However, mutation of C592 or C608 did impact stability and trafficking of ABCG2 to the cell membrane (Henriksen et al., 2005). Mutation of N596 prevents glycosylation which leads to increased ubiquitin-mediated proteasomal proteolysis (Nakagawa et al., 2009).

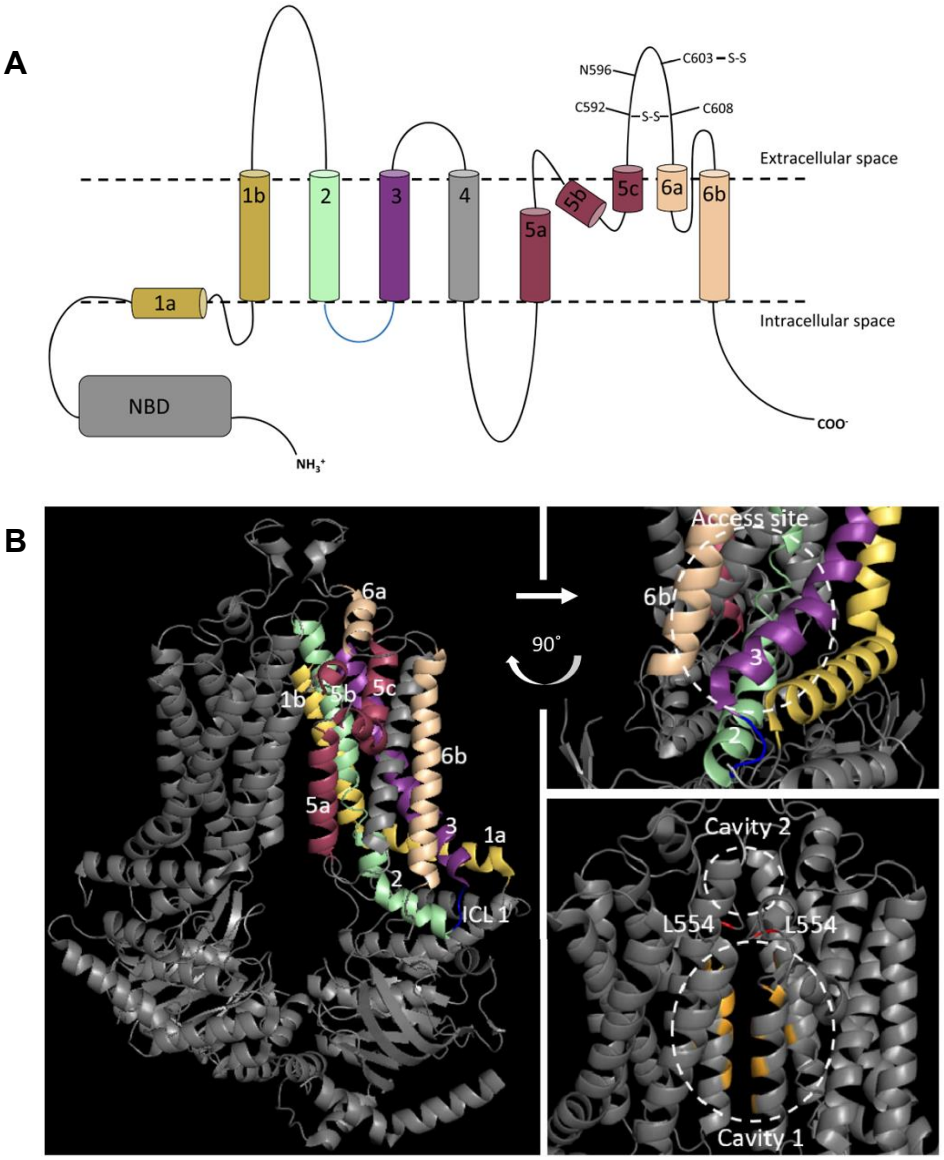


Figure 1.3 ABCG2 topology and structure. (A) Schematic representation of the topology of ABCG2 within the cell membrane. The transmembrane helices 1-6 (TM1-TM6) are shown in different colours. Key residues for intramolecular (C592 and C608) and intermolecular (C603) disulfide bonds and glycosylation (N596) are shown. **(B)** Cartoon representation of the ABCG2 dimer from PDB 6ETI (Jackson et al., 2018). TM1-6 are the same colours as in A. Rotation by 90° gives a clearer view of the "access site" which is surrounded by TM2, TM3 and TM6b. The bottom right panel shows the "leucine plug" (L554, red), cavity 1 and cavity 2. Figure taken from (Kapoor et al., 2018).

There are multiple binding sites for substrates, shown in Figure 1.3. The main site that the substrate binds to before being transported is cavity 1, which is located between the two ABCG2 monomers and surrounded by TM2 and TM5 (László et al.,

2016). Cavity 2 is on the extracellular side of ABCG2 and is the final binding site before the substrate is released into the extracellular space. Cavity 1 and cavity 2 are separated by a “leucine plug” which consists of L554 from each ABCG2 monomer and blocks movement of the substrate into cavity 2 before conformational changes occur (Taylor et al., 2017). Both these cavities are described by structural biology data (see below). A third binding site, referred to as the “access site” and predicted by mutagenesis studies and molecular modelling, is located between TM2, TM3 and TM6 and includes the residue R482 which has been linked to substrate specificity (Kapoor et al., 2018). The R482G mutation leads to broader substrate specificity, with drugs such as doxorubicin, daunorubicin and rhodamine 123 becoming transported when they otherwise would not (Ozvegy-Laczka et al., 2005, Tamura et al., 2007). Another substrate binding site (not shown) has also been predicted by molecular docking (László et al., 2016).

In recent years, cryo-EM structures have become available which has helped elucidate the structure and mechanism of ABCG2. The first was published by Taylor et al. (2017), which showed ABCG2 locked in an inward facing conformation by an inhibitory anti-ABCG2 antibody, 5D3, on the extracellular side. The NBDs did not interact to form the “ATP sandwich dimer” as no nucleotide was bound (as described in Figure 1.2) but were still in contact through a novel ABCG-family specific NBD:NBD interaction surface, consistent with other inward facing structures (Manolaridis et al., 2018, Jackson et al., 2018). Manolaridis et al. (2018) solved two further structures of ABCG2: the E211Q mutant ($ABCG2_{E211Q}$) bound to either the transport substrate estrone 3-sulfate (E_1S) or to the catalytic substrate ATP. This helped shed light on the mechanism of transport by ABCG2. As seen in Figure 1.4, the E_1S-bound $ABCG2_{E211Q}$ adopts an open, inward facing conformation with the

substrate located in cavity 1. The ATP-bound ABCG2$_{E211Q}$ shows a closed, outward facing conformation, where cavity 1 is completely closed, forcing the substrate towards cavity 2 and then the substrate is released on the extracellular side. The NBDs in this structure do dimerise in the classical fashion (Figure 1.2) with ATP-bound at the interface (Manolaridis et al., 2018). This structure is consistent with the idea that conformational change and transport of substrate occurs on ATP binding and not on ATP hydrolysis, as previously predicted by radioligand binding studies (McDevitt et al., 2008). ATP hydrolysis likely resets the inward facing conformation (Manolaridis et al., 2018).

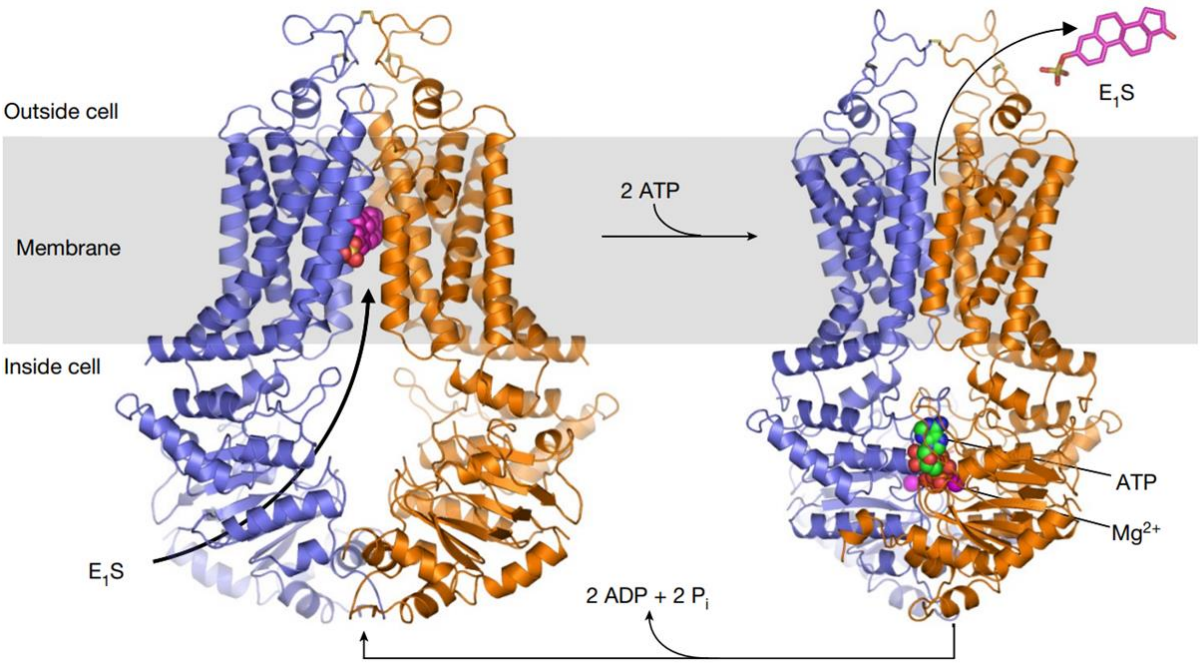


Figure 1.4 Mechanism of substrate transport by ABCG2. Cartoon representation of E_1S (substrate)-bound (left, PDB 6HCO) and ATP-bound (right, PDB 6HBU) ABCG2$_{E211Q}$. ABCG2$_{E211Q}$ monomers are displayed in blue and orange. Figure taken from Manolaridis et al. (2018).

1.2.4 Inhibitors

Inhibitors of ABCG2 prevent transport of substrates and therefore increase their cellular accumulation. This would also apply to drug molecules which are substrates, helping prevent multidrug resistance. There are currently no clinically available ABCG2 inhibitors but Ko143 is widely used in research. Ko143 is derived from a fungal toxin, fumitremorgin C (FTC), which is highly neurotoxic. Ko143 is more potent and less toxic, although still not clinically utilisable (Allen et al., 2002, Toyoda et al., 2019). Jackson et al. (2018) published the only current structures of ABCG2 bound to inhibitors: MZ29 (PDB 6ETI) and MB136 (PDB 6FEQ). MZ29 is a Ko143 derivative and binds to cavity 1, just as substrates do (Figure 1.5). Two MZ29 molecules bind between TM1b and TM2 of one monomer and TM5a of the other and vice versa. This was concordant with data indicating that a 2:1 inhibitor:ABCG2 molar ratio was required for full inhibition of ATPase activity (Jackson et al., 2018). The larger MB136 binds in the same location but only one molecule fits in cavity 1 and a 1:1 ratio is sufficient for maximum inhibition (Jackson et al., 2018). Since stable electron densities were found for these inhibitors in absence of the stabilizing antibody, 5D3, transport of these molecules is minimal. Binding of other inhibitors, such as febuxostat and elacridar, has not been structurally determined so it is unclear whether they inhibit allosterically (i.e. at a site distinct from cavity 1) or orthosterically (i.e. at cavity 1)(Toyoda et al., 2019).

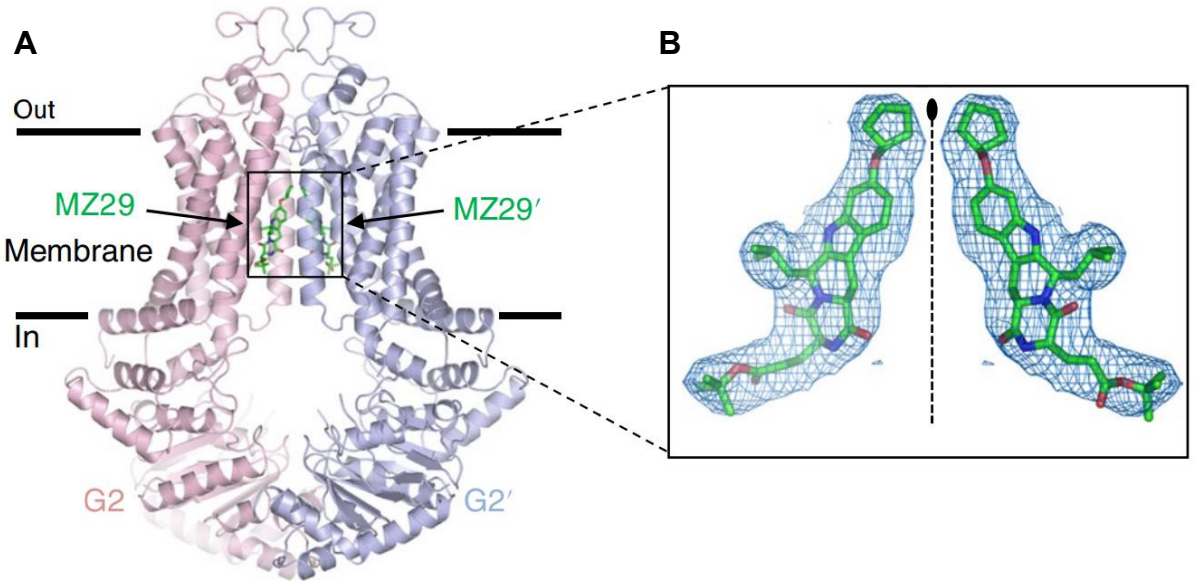


Figure 1.5 Structure of inhibitor-bound ABCG2. (A) Cartoon representation of the ABCG2 homodimer from PDB 6ETI. MZ29 (green sticks), a Ko143 derivative, is bound between the two ABCG2 monomers (pink and blue) in cavity 1. (B) Rotated (45°) view of the bound MZ29 molecules (green sticks) with their electron microscopy density (blue). The dotted line represents the two-fold symmetry axis. Both A and B were retrieved from Jackson et al. (2018).

1.3 Project aims

Distinguishing the features of ABCG2 inhibitor binding from features of ABCG2 substrate binding would be an important step in tackling multidrug resistance. The aim of this project is to use the cryo-EM structures of inhibitor-bound ABCG2 to identify key residues involved in inhibitor binding. Then relevant mutations of ABCG2 will be made and assessed for whether the mutations prevent inhibition or even transforms the inhibitor into a substrate.

Chapter 2 Materials and Methods

2.1 Materials and reagents

All molecular biology reagents were purchased from New England Biolabs (NEB, Hitchin, UK) and Promega (Southampton, UK), unless stated otherwise. The primers for site directed mutagenesis were ordered from Sigma-Aldrich (Poole, UK). All other cell culture materials and reagents, except Zeocin™ (Invitrogen) and polyethyleneimine (PEI, Polyscience Inc.), were obtained from Sigma-Aldrich or Thermo Fisher Scientific unless otherwise mentioned.

2.2 Molecular biology

2.2.1 Site-directed mutagenesis

Mutations were introduced into a WT-ABCG2 expression vector by performing mutagenic polymerase chain reactions (PCR). These were set up using mutagenic primers shown in Table 2.1 (0.5 µM), dNTPs (0.2 mM), pcDNA™3.1/Zeo(+)_SNAP_ABCG2[1] template (50ng), Pfu polymerase (1.0-1.5 U), and polymerase buffer made up to 50 µL. Denaturation, annealing and elongation steps were performed using a SensoQuest Thermocycler and cycled 18 times at the temperatures shown in Table 2.2 (with the annealing temperature varied until successful amplification was observed). Presence of amplified PCR product was confirmed by agarose gel electrophoresis. PCR products were mixed with Gel

[1] Zeo(+) refers to presence of a Zeocin™ resistance gene (BleoR). SNAP refers to the SNAP-tag® which is a modified alkylguanine DNA alkyltransferase which can specifically and covalently bind to benzyl guanine fluorophores as a way of tagging the protein of interest (Tirat et al., 2006). However, labelling of ABCG2 was not used in this project.

Loading Dye and loaded onto a 1% (w/v) agarose gel which was prepared in TBE buffer (10.8% (w/v) Tris, 5.5% (w/v) boric acid, 20 mM Na_2EDTA) and contained ~1 µg/mL ethidium bromide. The resolved gels were then viewed under UV light (using a Syngene GeneGenius gel imaging system).

Table 2.1 Forward primers used in the generation of the mutant ABCG2 constructs. Codon changes that result in an amino acid change are highlighted in yellow, uppercase letters indicate bases that are different from the template and lowercase letters identical to the template. The complementary reverse primers are not shown but have identical length, T_m and GC**.**

Name	Sequence (5'-3')	Length (bp)	T_m (ºC)	GC (%)
T435A	cttcttcctgacgGcTaaccagtgtttc	28	68.95	50.0
N436A	cttcctgacgaccGCccagtgtttc	25	71.82	60.0
F439A	gaccaaccagtgtGCcagcagtgtttc	27	71.78	55.56
S440W	caaccagtgtttTTgGagtgtttcagcc	28	70.13	46.43
M549E	ctgttttgtgttCatgGAgatttttcaggtctg	34	72.81	38.24
A397S/V401A	gcctctataTctcagatcattgCcacagtcgtactg	36	74.46	47.22
L539A	ctgtagcaacaGCtctTatgaccatctg	28	65.17	46.43
L405A	gtcacagtcgtaGCgggactggttatag	28	67.79	53.56
I543A/V546A	cttctcatgaccGCctgttttgCCtttatgatgatt	36	77.65	41.67

Table 2.2 PCR thermocycling parameters. Steps 2-4 were cycled 18 times before proceeding to step 5, as demonstrated by the arrow. Annealing temperatures were varied (50-65 °C) if these standard parameters were unsuccessful.

Step	Temperature (°C)	Time (min)	Number of cycles
1. Heat	95	1	1
2. Denature	95	1	18
3. Anneal	55	1	18
4. Extend	72	12	18
5. Extend	72	10	1
6. Cool	10	∞	1

2.2.2 *DpnI* digest and transformation

To remove template DNA, PCR products were digested with *DpnI* (20 units) for 90 minutes at 37 °C. The enzyme was then deactivated by incubating at 80 °C for 20 min. *DpnI* digested-PCR product was transformed into DH5α competent *E. coli*. First 100 µL of DH5α cells were thawed on ice and 5 µL of the *DpnI* digested PCR product was added. The competent cell mixture was then incubated on ice for 1 hour, heat

shocked at 42 °C for 1 min, and then placed back on ice for 2 min. The cells were then supplemented with 250 µL Luria-Bertani (LB) medium (1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract) and shaken at 37 °C for 1 hour. All 350 µL was subsequently spread onto LB-agar plates (1.5% (w/v) agar, 100 µg/mL ampicillin) and incubated overnight at 37 °C.

2.2.3 Plasmid preparation

Single colonies from LB-agar plates were grown overnight at 37 °C with shaking, after being picked and inoculated into 5 mL of ampicillin-supplemented (100 µg/mL) LB medium. For long-term storage, the resultant bacterial cultures were stored at -80 °C as glycerol stocks (500 µL of bacterial culture, 500 µL of 30% (v/v) glycerol). The remainder of the cultures were centrifuged for 5 mins at ~3000 x g and the supernatants were discarded. Plasmid DNA was then extracted from the resulting bacterial pellet using a NucleoSpin® Plasmid kit (Macherey-Nagel). The manufacturer's protocol was followed except instead of centrifuging at 11,000 x g, the samples were centrifuged at 13,000 x g.

DNA plasmid concentration and purity were determined using the Nanodrop 2000® (Thermo Fisher Scientific). The purity was confirmed by the A260/A280 ratio, where a value of >1.7 was deemed acceptable for future transfection.

2.2.4 Sequencing

To confirm presence of the desired experimental mutations and absence of other mutations, the purified DNA underwent Sanger sequencing. The mutant constructs were sent to Source Bioscience (Nottingham, UK) along with the primers shown in Table 2.3. The quality of the chromatograms was analysed in Chromas

(Technelysium Pty Ltd) and the sequences were aligned with the template sequence using BLAST (NCBI) to ensure only the desired mutation was incorporated plasmid DNA.

Table 2.3 Sequencing primers. A series of primers were used to sequence the full SNAP-ABCG2 coding region (bases 1026-3543 of the pcDNA™3.1/Zeo(+)_SNAP_ABCG2 plasmid). "Region sequenced" refers to the minimum region where sequencing data was of a high quality for all mutants.

Primer name	Sequence (5'-3')	Region sequenced (bp)
T7 promoter (F)	TAATACGACTCACTATAGGG	978-1992
SeqF1	CACAGGTGGAGGCAAATCTT	1935-2852
SeqF2	GCAGGGACGAACAATCATCT	2361-3310
Seq482	AACTCTTTGTGGTAGA	3028-3749

2.3 Cell culture

2.3.1 Maintenance of cell cultures

HEK293T cells were grown in T25 (25 cm^2) flasks at 37 °C, 5% CO_2 in Dulbecco's modified eagle medium (DMEM, 4500 mg/L glucose, L-glutamine, sodium pyruvate, and sodium bicarbonate) supplemented with 10% (v/v) foetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin (P/S, 100 U/mL penicillin, 100 µg/mL streptomycin). Once cells reached 70-90% confluency based on visual inspection, the cultures were passaged (typically twice a week). Media was removed, cells were washed once with pre-warmed phosphate buffered saline (PBS) and incubated with 0.5 mL trypsin/EDTA for 1-3 min. The trypsin/EDTA was then quenched with 4.5 mL of DMEM and cells were then detached by repeated pipetting, before being centrifuged 225-250 x g for 5 min. The cell pellets were then resuspended in DMEM and reseeded at a typical dilution of 1/10.

2.3.2 Transfection

Cells were seeded at 1.25 x 10^5 cell/mL (determined using a haemocytometer) into a 6-well plate, ~24 hours prior to transfection. Two hours before transfecting, the media was replaced with 5% serum media (DMEM, 5% FBS, 0.5% P/S). Cells were transfected using linear polyethyleneimine (PEI) at a molar PEI nitrogen:DNA phosphorus ratio of 15:1. Transfection mixtures were made by the addition of 9 µL of PEI (from a 10 mM working stock solution) to 2 µg DNA, as described by Cox et al. (2018), and were left no longer than 5 minutes before adding 100 µL of media and dropwise addition to the HEK293T cells. 24 hours later media was replaced with 10% FBS-supplemented media.

2.3.3 Zeocin™ selection

Transfected cells were transferred to T25 flasks by trypsinization (described in 2.2.1 and scaled down appropriately). Flasks contained 5 mL DMEM (10% FBS, 1% P/S) and approximately 5 hours later, Zeocin™ was added to total concentration of 200 µg/mL. Media was changed every 2-3 days and Zeocin™ concentration was maintained at 200 µg/mL for several weeks until growth of Zeocin™ resistant colonies was observed. Transfected cells expressing sfGFP-ABCG2 were employed as a control for Zeocin™ selection. sfGFP tagged proteins were chosen because protein expression (fluorescence) could be monitored under an epifluorescence microscope without addition of fluorescent reagents. This helped confirm when the other cells lines were stably transfected. At this point, Zeocin™ concentration was dropped to 40 µg/mL.

2.3.4 Epifluorescence microscopy

Expression of GFP-ABCG2 in pcDNA™3.1/Zeo(+)_sfGFP_ABCG2 transfected HEK293T cells was monitored with epifluorescence microscopy. Images were obtained using the Zeiss Axiovert S100 microscope and Zeiss AxioCam MRm monochrome digital camera and processed with AxioVision version 4.8.2 SP3.

2.3.5 Long-term storage

For long-term storage, 80% of cells in a T25 flask were resuspended in 4 mL ice cold freezing medium (10% (v/v) DMSO in FBS) and aliquoted into 1 mL cryovials. The other 20% were maintained as described in 2.2.1. The cryovials were then frozen slowly (~1 °C/min) by storage in a precooled (4 °C) Mr Frosty™ (containing isopropanol) in a -80 °C freezer. Frozen cryovials were transferred to liquid nitrogen for longer term storage. When needed, cells were rapidly thawed by pipetting up and down with warm media, centrifuged at 225-250 x g, resuspended in 5 mL DMEM (10% FBS, 1 %P/S) and maintained in a T25 flask as described in 2.2.1.

2.4 SDS-PAGE and western blotting

2.4.1 Cell harvesting and lysis

Cell culture monolayers were washed once with PBS, detached with Trypsin/EDTA and centrifuged at 235 x g for 5 minutes (as described in 2.2.1). The pellet was then washed once with PBS and stored at -80 °C if not required immediately. Pellets were resuspended in 250 µL ice cold lysis buffer consisting PBS supplemented with 10% (v/v) glycerol and EDTA-free protease inhibitor cocktail III (Calbiochem, 1:200 dilution). Cells were then sonicated at 40% output for 4 x 5 seconds, being stored on ice for at least 2 minutes between bursts.

2.4.2 Protein concentration determination

To ensure equal protein loading in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis, a modified Lowry assay was performed. Using a Bio-Rad DC protein assay kit and a standard curve of 0-10 µg bovine serum albumin (BSA), total protein concentration of the cell lysates was determined. All samples and the standard curve were analysed in duplicate.

2.4.3 Sample preparation and SDS-PAGE

SDS-PAGE was performed as described by Laemmli (1970). Equal quantities of cell lysate (typically 20-100 µg) were incubated with protein loading buffer (50 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 100 mM 2-mercaptoethanol) at 37 °C for 30 minutes. Resolving gels (10% (w/v) polyacrylamide, 0.175% (w/v) SDS, 0.15% (w/v) ammonium persulfate (APS), 0.05% (v/v) N,N,N',N'-tetramethylethane-1,2-diamine (TEMED), 375 mM Tris base pH 8.8) and stacking gels (4% (w/v) polyacrylamide, 0.175% (w/v) SDS, 0.15% (w/v) APS, 0.06% (v/v) TEMED, 125 mM Tris base pH 6.8) were prepared before being placed in an electrophoresis tank filled with protein running buffer (25 mM Tris Base, 192 mM glycine, 3.5 mM SDS). Samples were loaded alongside a molecular weight marker (Invitrogen™ SeeBlue™ Plus2 Pre-stained Protein Standard) and electrophoresed at constant current (30 mA) until loading dye had fully eluted. Gels were then either used in a western blot (section 2.4.4) or stained with InstantBlue™ by rocking at room temperature for 2 hours. An ABCG2-positive sample was run on every gel as a control for the western blots.

2.4.4 Western blotting

Proteins were transferred from the polyacrylamide gel (section 2.4.3) to a nitrocellulose membrane by electrophoresis (200 mA, 2 hours, 4 °C) in transfer buffer (25 mM Tris base, 192 mM glycine, 20% (v/v) methanol), first described by Towbin et al. (1979). Transient staining of the nitrocellulose with Ponceau S solution (0.1% (w/v) Ponceau S, 1% (v/v) acetic acid) allowed confirmation of effective transfer of proteins. Membranes were washed with PBS-T (0.1% (v/v) Tween-20 in PBS) for 5 minutes, then incubated with non-fat milk (5% (w/v) in PBS-T) at room temperature for 1 hour to prevent antibody from binding to non-specific sites. The blots were subsequently incubated with the primary antibody (anti-ABCG2 antibody BXP-21, 1:2000 dilution in non-fat milk) overnight at 4 °C. Then the blots were washed with PBS-T several times over a period of 15-20 minutes to remove any unbound primary antibody and then incubated with the secondary antibody (rabbit anti-mouse horseradish peroxidase, 1:5000 dilution in non-fat milk) at room temperature for 1 hour. The 15-20-minute washes were then repeated before incubating with the chemiluminescence substrate (SuperSignal™ West Pico PLUS, Thermo Fisher Scientific) for 1 minute and imaging with the LAS-3000 mini (Fujifilm).

2.5 Flow cytometry

2.5.1 Cell surface expression

ABCG2-transfected and untransfected HEK293T cells were seeded at 1 x 10^6 cells/mL in FACS buffer (0.2-1.0% fatty acid-free BSA in PBS or Hank's balanced salt solution, HBSS) and incubated for 30 mins on ice with 5D3 antibody (1.82-5.00 µg/mL, Millipore), isotype control (1 µg/mL anti-TRP1, Santa Cruz Biotechnology or 3 µg/mL mouse IgG isotype control, Invitrogen) or nothing (negative control). Cells

were then centrifuged at 350 x g for 5 min at 4 °C and the supernatant was discarded. Cell pellets were washed twice by resuspending in FACS buffer and spun down at 350 x g for 5 min at 4 °C, twice. Then, the 5D3 and isotype control cells were incubated with AlexaFluor647 (5-10 µg/mL goat anti-Mouse IgG (H+L) Alexa Fluor 647, Invitrogen) in 1 mL FACS buffer for 1 hour on ice. Cells were then centrifuged at 350 x g for 5 min at 4 °C, the supernatant was discarded and washed twice. Finally, cells were resuspended in FACS buffer and fluorescence was measured by flow cytometry using the Beckman Coulter Astrios EQ Cell Sorter (channel 640-671/30).

Data was initially analysed using Kaluza 2.1. Firstly, side scatter height (channel 488-SSC) was plotted against forward scatter height (channel 488-FSC1) and was gated to exclude debris from live cells. Then side scatter height was plotted against side scatter area and gated to allow monodispersed cells to be separated from doublets. Then a histogram of AlexaFluor647 fluorescence vs number of cells was plotted and the gate placed at the edge of the isotype control peak was used to determine percentage of positive cells i.e. percentage of cells with more fluorescence with 5D3 than with the isotype control.

2.5.2 Transport assay

ABCG2-transfected and untransfected HEK293T cells were seeded at 1 x 10^6 cells/mL in phenol red-free DMEM and incubated at 37 °C, 5% CO_2 for 1 hour with either DMSO (0.1% (v/v)) or fluorescent derivatives of Ko143 (2 µM, Ko143-Cy5 or Ko143-X-BY630; synthesized by Sarah Mistry, School of Pharmacy, University of Nottingham). After the incubation, cells were kept on ice. Cells were then centrifuged at 350 x g for 5 min at 4 °C and the supernatant was discarded. Cell pellets were then resuspended in phenol red-free DMEM and spun down again at 350 x g for 5

min at 4 °C. Finally, cells were resuspended in phenol red-free DMEM and fluorescence was measured by flow cytometry using the Beckman Coulter Astrios EQ Cell Sorter (channel 640-671/30). Flow cytometry is capable of simultaneously measuring ABCG2 function and expression for each cell in suspension. This allows the gating out of lower ABCG2-expressing cells when looking at the transport data, however, it was not possible in this project.

Data was gated for monodispersity as in 2.5.1. Then, the median fluorescence of the DMSO control was subtracted from the median fluorescence of the fluorescent Ko143-treated cells. This removes background fluorescence giving a value for ΔMFI (change in median fluorescence intensity).

$$\Delta MFI = MFI_{Fluorescent\ Ko143} - MFI_{DMSO}$$

2.6 Monolayer transport assay

A 96-well plate was pre-treated with poly-L-lysine (10 µg/mL, 100 µL in each well) for 1 hour, then aspirated and left to dry. Subsequently, 3 x 10^4 HEK293S cells (HEK293 cells adapted for growth in suspension (Lin et al., 2014)), expressing WT-ABCG2 or not, were seeded. 48 hours later, media was removed and cells were incubated at 37 °C with either DMSO; substrate (Hoechst 33342) in the presence or absence of inhibitor (Ko143 or a fluorescent Ko143 derivative); or a fluorescent Ko143 derivative in the presence or absence of Ko143. Concentrations are shown in Table 2.4. After 45 min, media was removed and cells were incubated with either DMSO (1:1000), fluorescent Ko143 derivative (1 µM), or Ko143 (1 µM) as shown in Table 2.4. Then, after another 45 min, media was replaced with HBSS. Hoechst 33342 fluorescence (excitation 350 ± 10 nm, emission 460 ± 20 nm) and Ko143-X-BY630 fluorescence (excitation 620 ± 10 nm, emission 660 ± 20 nm) or Ko143-Cy5 fluorescence

(excitation 610 ± 30 nm, emission 675 ± 50 nm) was measured with the CLARIOstar microplate reader (BMG Labtech). The CLARIOstar microplate reader measures average fluorescence of each well so it is not possible to separate low ABCG2-expressing cells from higher expressing cells. All dilutions were made with HBSS and DMSO concentration was the same in each well. Fluorescent derivatives used above were Ko143-X-BY630 or Ko143-Cy5. Data was collected in collaboration with James Mitchell-White (Kerr Lab, School of Life Sciences, University of Nottingham). Background fluorescence was considered by subtracting the average DMSO fluorescence value from each datapoint.

$$Fluorescence = Fluorescence_{Raw} - Average\ Fluorescence_{DMSO}$$

Where $Fluorescence_{Raw}$ is the Hoechst 33342 or fluorescent Ko143 derivative fluorescence for an individual datapoint and $Average\ Fluorescence_{DMSO}$ is the mean fluorescence value of DMSO treated cells for each plate and cell line.

Table 2.4 Monolayer transport assay incubation steps First incubation step involves substrate, inhibitors or DMSO control diluted with HBSS. The corresponding second incubation step is also shown. For Ko143-Cy5 experiments, Ko143-Cy5 was used in place of Ko143-X-BY630.

First incubation	Second Incubation
DMSO	DMSO
Hoechst 33342 (3 μM)	DMSO
Hoechst 33342 (3 μM) + Ko143-X-BY630 (1 μM)	Ko143-X-BY630 (1 μM)
Hoechst 33342 (3 μM) + Ko143 (1 μM)	Ko143 (1 μM)
Ko143-X-BY630 (1 μM)	DMSO
Ko143-X-BY630 (1 μM) + Ko143 (1 μM)	Ko143 (1 μM)

2.7 Data analysis

All data were analysed using GraphPad Prism 9.1.0. Multiple data sets were compared with WT and untransfected HEK293T cells using a one-way ANOVA with a Dunnett's multiple comparisons test. All experiments were repeated a minimum of 3 independent occasions for statistical analysis and figure legends confirm the number of technical repeats (n). The monolayer transport assay data was analysed with a two-way ANOVA with Tukey's multiple comparisons test.

Chapter 3 Results

3.1 Hypothesis formation

The hypothesis studied in this project was based around the question, “what makes an inhibitor an inhibitor and not a substrate that is transported by ABCG2?”. This could potentially be explained by differences in affinity, with inhibitors having much higher affinity or potency than substrates, whether that be measured by dissociation constants or by IC_{50} values, as suggested by Jackson et al. (2018). Further investigation of the literature, provided some evidence for this, with inhibitors having between 3-fold and ~2000-fold higher affinity than substrates (Table 3.1) The only apparent exceptions to this are TLC-S and sulfasalazine. Sulfasalazine has been used as a substrate and an inhibitor in various research (Miyata et al., 2016, Karlsson et al., 2010). TLC-S (taurolithocholate sulfate) is a bile acid which at certain concentrations can disrupt the cell membrane which could explain its slight inhibitory effect but low affinity for ABCG2 (Chiang, 2003). The apparent affinity values, shown in Table 3.1, are the concentration at which either half maximal transport occurred or half maximal inhibition was achieved (K_m or IC_{50}). These values cannot be directly compared and IC_{50} values will vary depending on the specific conditions used in the assay but can act as a guide to how strongly these compounds bind to ABCG2. Jackson et al. (2018) also found a ~3,000-fold difference in affinity between their fluorescent Ko143 derivative (inhibitor) and the substrate E_1S.

Table 3.1 Apparent affinities of substrates and inhibitors

Compound name	Substrate or Inhibitor	Apparent affinity (μM)	Reference
Fumitremorgin C	Inhibitor	0.731 ± 0.092	(Ochoa-Puentes et al., 2013)
Ko143	Inhibitor	0.128 ± 0.017	(Kohler and Wiese, 2015)
Benzbromarone	Inhibitor	0.20	(Miyata et al., 2016)
Topiroxostat	Inhibitor	0.18	(Miyata et al., 2016)
Febuxostat	Inhibitor	0.027	(Miyata et al., 2016)
Sulfasalazine	Substrate/Inhibitor	0.6	(Karlsson et al., 2010)
TLC-S	Inhibitor	37	(Suzuki et al., 2003)
Rosuvastatin	Substrate	2.3	(Miyata et al., 2016)
E_1S	Substrate	16.6 ± 3.4	(Suzuki et al., 2003)
Mitoxantrone	Substrate	61	(Suzuki et al., 2003)
DHEAS	Substrate	55	(Suzuki et al., 2003)
4-MUS	Substrate	12.9 ± 2.1	(Suzuki et al., 2003)
E3040S	Substrate	26.9 ± 4.0	(Suzuki et al., 2003)
SN-38	Substrate	4.0	(Nakatomi et al., 2001)
SN-38-glucuronide	Substrate	26	(Nakatomi et al., 2001)

The difference in affinity can be rationalised by the cryo-EM structures determined by Manolaridis et al (2018). The substrate-bound (E_1S), inward facing ABCG2 structure (PDB 6HCO) showed that F439 from opposite monomers are 8 Å apart and interacting with the bound substrate. In ATP-bound, outward facing structure (PDB 6HBU) these residues are only 3.5 Å apart (Figure 3.1). This means the substrate must be released from cavity 1 before it completely collapses, suggesting transient conformational changes which push out the substrate, resembling a peristaltic motion (Manolaridis et al., 2018). The resulting hypothesis is that inhibitors have a higher affinity for cavity 1 which means they are less likely to be released to allow the

transient conformational changes to occur. *In other words, high affinity inhibitors are proposed to lock ABCG2 in the inward facing conformation.*

One extension of this theory is that there would be an inverse relationship between binding affinity and maximal transport, suggested by Manolaridis et al. (2018). This group removed the hydrogen bond potential in a T435A mutant causing an increase in E_1S transport. This could indicate that reducing affinity for a compound increases substrate character (i.e. increased transport) and increasing affinity increases inhibitor character (less transport, more inhibition).

If inhibitors have high affinity and substrates have lower affinity, will reducing the affinity of an inhibitor cause it to become a substrate which is transported? To test this hypothesis a series of experimental mutants were made with the goal of reducing affinity for the most commonly used inhibitor Ko143.

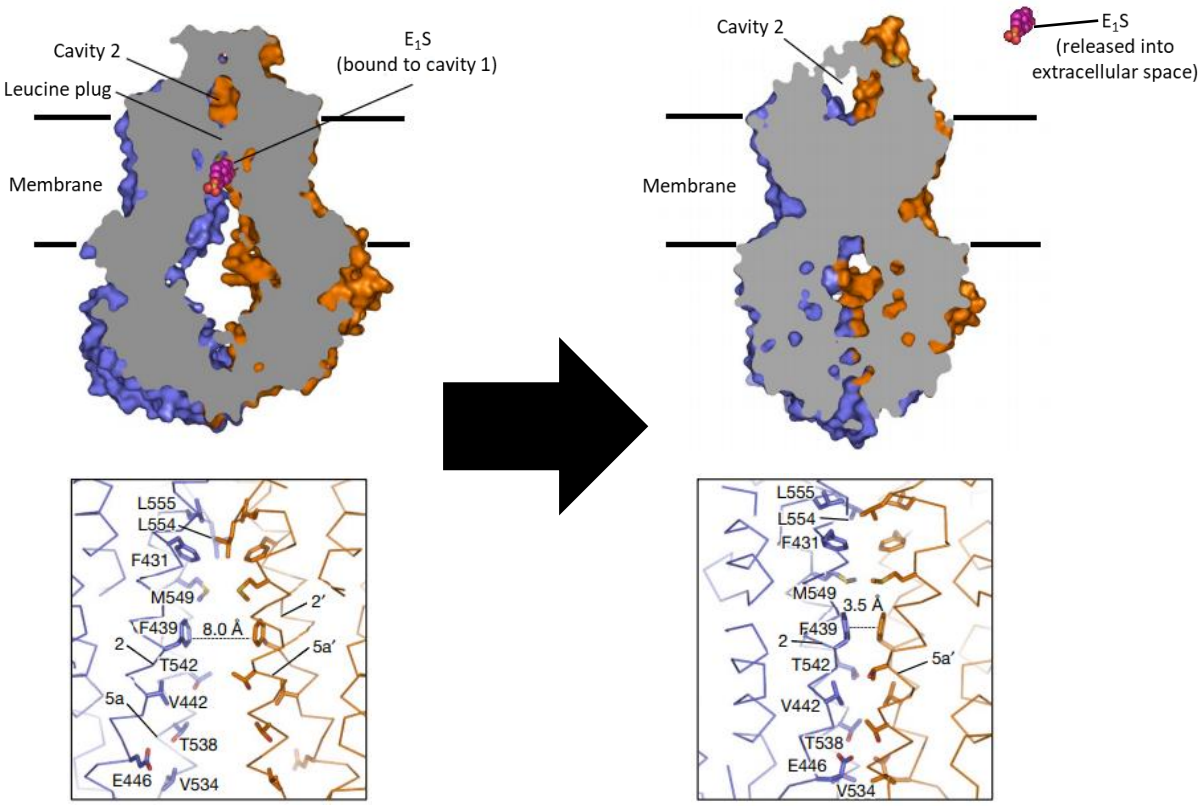


Figure 3.1 Cavity 1 completely collapses to promote substrate transport. When the substrate E_1S is bound in cavity 1 of human ABCG2 (top left, PDB 6HCO), the F439 residues from each monomer are 8.0 Å apart (bottom left). ATP binding triggers conformational changes which lead to E_1S being extruded into the extracellular space (top right, PDB 6HBU) and cavity 1 collapses so that F439 residues are only 3.5 Å apart (bottom right). Cross sections (top left and top right) and stick models (bottom left and bottom right) of each monomer are shown in orange and blue. The leucine plug residue is L554. Figure adapted from Manolaridis et al. (2018).

3.2 Mutant design

In order to decide which experimental mutations to make, the structure of ABCG2 was visualised in PyMOL. This allowed determination of residues that are involved in the binding of Ko143 and provided information for the design of mutants that would have a reduced affinity for Ko143. The structure used was the cryo-EM structure published by Jackson et al. (2018) of ABCG2 bound to MZ29, a derivative of Ko143. MZ29 and Ko143 only differ in one place: the methoxy group of Ko143 is replaced by an O-cyclopentyl group (Figure 3.2). First, all residues with atoms within 4.0 Å of

MZ29 were identified; a total of fifteen were found. Three of these residues (F431, F432 and L555, shown in white in Figure 3.3 B and C) were excluded since they only interacted with the O-cyclopentyl group of MZ29 and would be too far away (4.7-5.9 Å) to interact with the carbon of the methoxy group in Ko143. Also, Manolaridis et al. (2018) found that no functional protein was expressed with L555A because it is likely to have structural importance, so mutations to different residues could have the same issue. The remaining residues are shown in Figure 3.3 A, where their interactions with specific parts of Ko143 are also shown.

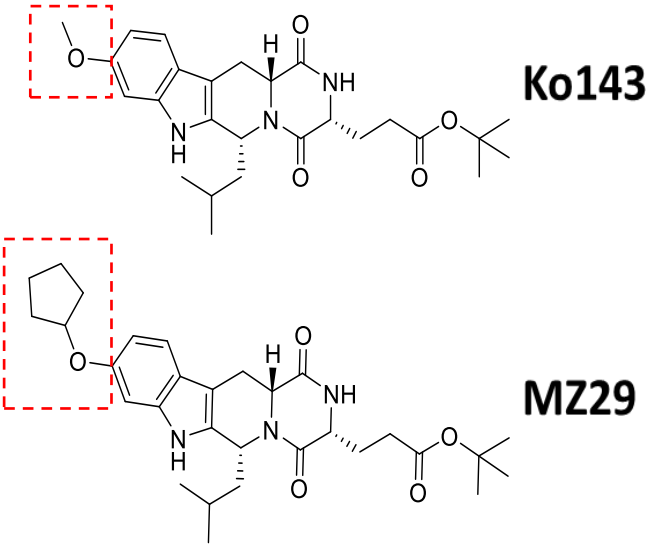


Figure 3.2 Comparison of the structures of Ko143 and MZ29. The red dashed squares highlight the difference between Ko143 and MZ29: a methoxy group and O-cyclopentyl group respectively. Structures are from Jackson et al. (2018).

In an attempt to narrow down the important residues further, the human ABCG2 protein sequence was aligned with ABCG2 sequences from other mammals (rabbit, cow and mouse) where Ko143 has been shown to be an inhibitor (Weidner et al., 2015, Manzini et al., 2017, Wei et al., 2012, Halwachs et al., 2016). All residues with atoms within 4.0 Å of MZ29 were fully conserved (Figure 3.3 D-F), in fact most residues in ABCG2 are conserved (81.61-86.28% sequence identity compared with

human ABCG2). Therefore, all highlighted residues could play a role in binding, and none can be ruled out at this stage based off the sequence alignments alone.

3.2.1 A397S/V401A/L539A – Mut1

A397, V401 and L539 interact with the *tert*-Butyl ester (blue in Figure 3.3 A) via van der Waals forces. Mutating these residues to reduce hydrophobicity could reduce the affinity of ABCG2 for Ko143. This is supported by the research performed by Weidner et al. (2015) where the acid metabolite of Ko143, lacking the *tert*-Butyl ester, has no inhibitory effect on mitoxantrone transport by ABCG2. V401A and L539A would lower affinity for Ko143 since mutation to alanine would introduce a smaller hydrophobic group which would form fewer or weaker van der Waals interactions. However, this approach could not be applied to A397 which cannot be mutated to anything smaller, besides glycine which would add flexibility to α-helix 1b (Figure 3.3 B, C and D), potentially causing unwanted effects to the overall protein structure (Högel et al., 2018). For this reason and to further decrease affinity, the mutation to the polar residue serine was introduced in this triple mutant. Of the three residues interacting with the *tert*-Butyl ester, A397S is the furthest away from the oxygens and is therefore the least likely to hydrogen bond but could still cause a hydrophobic/hydrophilic repulsion that would result in a decrease in affinity. Making the triple mutant, A397S/V401A/L539A, will ensure the hydrophobicity is sufficiently reduced to have an impact on the affinity for Ko143's *tert*-Butyl ester. This mutant will be referred to as Mut1 throughout this thesis.

3.2.2 L405A/I543A/V546A – Mut2

The isobutyl group, highlighted by the green circle in Figure 3.3 A, forms van der Waals interactions with V401, L405, I543 and V546. Another triple mutant was designed to target these hydrophobic interactions: L405A/I543A/V546A (Mut2). Cox et al. (2018) made the single mutants L405A and I543A, both of which showed a significant reduction in mitoxantrone and pheophorbide A transport. It is, therefore, feasible these residues will also play a role in the binding of inhibitors. V546A, however, compared to WT had no change in E_1S transport but did double the EC_{50} of E_1S-induced ATPase activity (Manolaridis et al., 2018). This means the V546A mutant reduces the affinity for the substrate E_1S but perhaps not enough for reduced transport. By combining V546A with L405A and I543A the hydrophobicity will be lowered to a greater extent which will have a larger effect on reducing the affinity, when applied to Ko143.

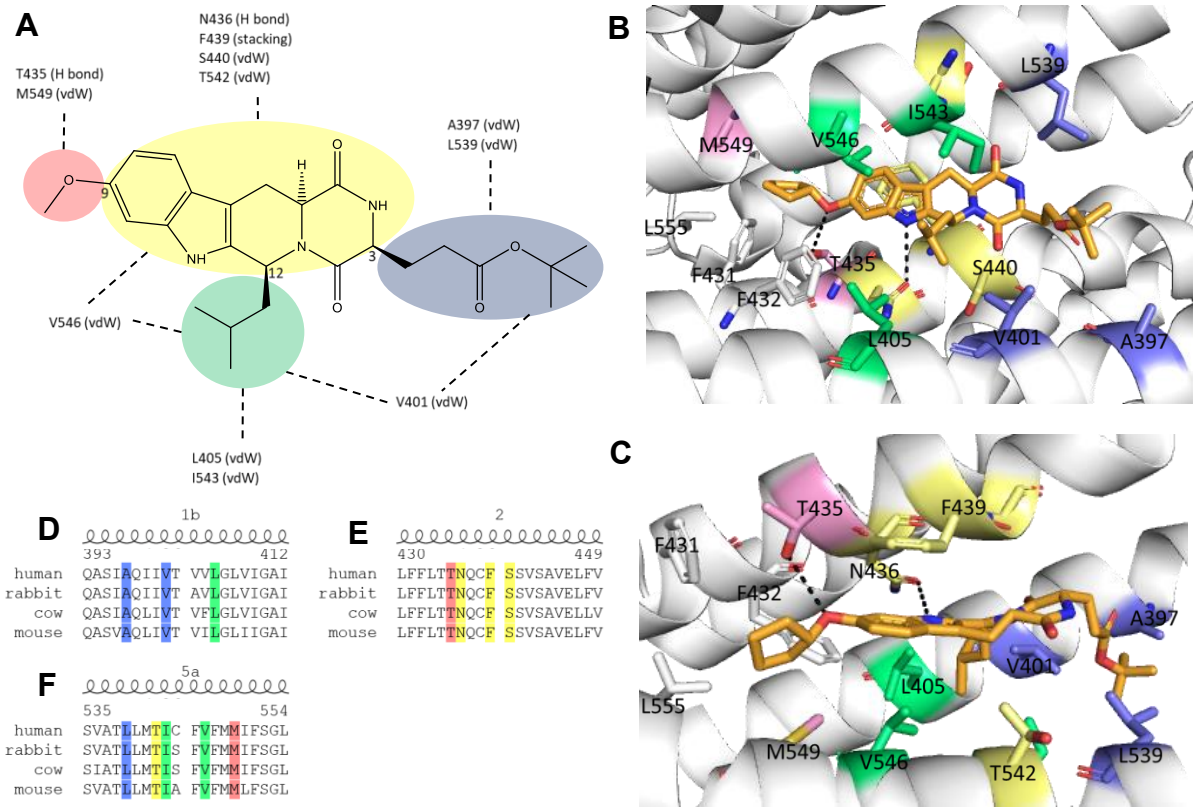


Figure 3.3 ABCG2 residues interacting with Ko143. Schematic showing ABCG2 residues interacting with Ko143. Four chemical features of Ko143 are involved in interactions: methoxy (pink), polycyclic core (yellow), *tert*-Butyl ester (blue), isobutyl (green). Carbons 3, 9 and 12 are labelled (Jackson et al., 2018) **(B and C)** MZ29-bound ABCG2 with interacting residues shown as sticks (within 4.0 Å). Polar interactions (hydrogen bonds) shown as black dotted lines. Residues interacting with one feature of MZ29 are coloured as in A. V401 (blue) is coloured to match the other residues interacting with the *tert*-Butyl ester (triple mutant) but it also interacts with the isobutyl group. V546 (green) is coloured in the same manner but also interacts with the polycyclic core. F431, F432 and L555 (white) interact with *O*-cyclopentyl group of MZ29 but not Ko143. Some residues from TM1b and TM5a (opposite monomers) are hidden for clarity. PDB 6ETI was used (Jackson et al., 2018) **(D-F)** Partial sequences from the alignment of ABCG2 from human, rabbit, cow and mouse. Shows conservation of residues interacting with Ko143 in TM1b (D), TM2 (E) and TM5a (F) (highlighted with same colours as in B and C).

3.2.3 T435A and M549E

Single mutants were made to target the affinity for the methoxy group of Ko143 (pink in Figure 3.3). This is because the methoxy group has been shown to be a key component of potent inhibitors, with Ko143 being four times more potent than its demethoxy analogue Ko134 when measuring inhibition of ABCG2-mediated mitoxantrone efflux (Allen et al., 2002). Demethoxy fumitremorgin C is also 10 times less potent than native fumitremorgin C (He et al., 1999, as cited in Allen et al., 2002) so small changes to the interactions with the methoxy will potentially have a greater effect than with other groups. This is supported by Jackson et al. (2018). They found that changing the substituent at C9 (the methoxy group, Figure 3.3 A) had a large influence on binding affinity and suggested that the removal of the T435 hydrogen bond causes a reduction in affinity. For these reasons, the mutation T435A was made which removes the hydrogen bond.

M549 also interacts with the methoxy group via van der Waals forces and has been shown to be important in substrate transport. In previous studies, M549A decreased mitoxantrone and pheophorbide A transport but not E_1S transport (Haider et al., 2015, Manolaridis et al., 2018). Jackson et al. (2018) found that the addition of hydrophilic groups to C9 of the polycyclic core of Ko143 caused the inhibitor to become inactive. They suggest this is due to the hydrophobicity at the bottom of cavity 1, nevertheless, it would be interesting to see the effect of mutating M549 to a glutamic acid which is a similar sized but hydrophilic residue. This mutation would remove the hydrophobic interactions without adding a hydrogen bond between the carboxyl and the methoxy (4.7 Å between 2 closest oxygen atoms). In addition, since

M549A had no effect on E_1S transport, a more drastic change (M549E) is more likely to have an effect on affinity for Ko143.

3.2.4 N436A, F439A and S440W

The polycyclic core of Ko143 (yellow in Figure 3.3 A) has 5 amino acids interacting with it: N436, F439, S440, T542 and V546. N436 hydrogen bonds via the carbonyl oxygen of the side chain with the NH of the indole ring in Ko143 (Figure 3.3 C). Mutating to an alanine removes this hydrogen bond and minimizes advantageous hydrophobic interactions with the polycyclic core. Previous mutation of N436A led to depleted E_1S transport activity and there was also no E_1S-induced ATPase activity suggesting it is important for substrate binding, therefore, the mutation N436A could reduce affinity for Ko143 (Manolaridis et al., 2018).

F439 forms stacking interactions with the benzene ring in the polycyclic core of Ko143, this is a stronger bond than van der Waals but weaker than hydrogen bonds so would be interesting to mutate. Removing the aromaticity of phenylalanine would eliminate this interaction. Mutating to any other amino acid, besides tryptophan or tyrosine, would do this but introduction of other interactions must also be avoided. This is why mutation to alanine was made: no hydrogen bonds are introduced and hydrophobic interactions are minimized. Previous data on the F439A mutant has shown that it had an effect on E_1S transport and E_1S-induced ATPase activity so could also have an effect on Ko143 affinity (Manolaridis et al., 2018).

S440 has been shown to be involved in mitoxantrone and pheophorbide A transport, with mutation to alanine causing a significant reduction in transport compared with the WT (Cox et al., 2018). However, the hydroxyl group of S440 is pointing away from MZ29 in the 6ETI structure (Figure 3.3 C) so would not form any polar

interactions with MZ29 nor Ko143, based on the assumption that they bind in the same manner. Perhaps the orientation of S440 is different when binding to substrates (although S440 also points away in the E_1S-bound structure, 6HCO) or the hydrophilicity of this residue contributes to creating an environment suitable for binding. Since S440A will not necessarily have a clear effect on the affinity for Ko143, a tryptophan mutation was made. The idea was to mutate to the largest amino acid to cause steric hindrance so that Ko143 can no longer sit in the pocket in a way that optimises the other interactions. This could have major ramifications on binding by causing too much steric hindrance or removing too many other interactions. On the other hand, interactions could be introduced including π-π stacking and hydrogen bonds with the NH of tryptophan.

T542 forms van der Waals interactions with the polycyclic core of Ko143, however, this residue was not mutated. Removing one van der Waals interaction would likely have very little effect on the total affinity especially since there are 4 other amino acids interacting with the polycyclic core, including a stronger hydrogen bond and stacking interaction. Furthermore, there was no literature to support its involvement in Ko143 or substrate binding. Since the aim of this project is to assess the importance of affinity for inhibitor function, there is no need to mutate this residue unnecessarily.

3.2.5 Summary

In summary, the mutants designed are: Mut1 (A397S/V401A/L539A), Mut2 (L405A/I543A/V546A), T435A, M549E, N436A, F439A and S440W. These mutations target different functional groups of Ko143 and are proposed to affect the affinity in different ways (Table 3.2). For example, Mut1 and Mut2 remove multiple weaker interactions, whereas T435A and N436A remove stronger hydrogen bonds. M549E

and S440W are mutating residues that are potentially less important for affinity, to more extreme mutations designed to disrupt binding. In future experiments, if the triple mutations had a functional impact, there would be value in assessing the contribution of each residue to Ko143 binding and transport. However, for this project the compounded effect of three minor changes in both Mut1 and Mut2 is more likely to noticeably impact Ko143 binding.

Table 3.2 ABCG2 mutants and their desired effect on Ko143 binding. Mut1 and Mut2 are A397S/V401A/L539A and L405/I543A/V546A triple mutants respectively.

Mutant	Ko143 functional group	Desired effect
T435A	Methoxy group	Removes hydrogen bond
M549E	Methoxy group	Removes hydrophobic interaction; Introduces hydrophilicity
N436A	Polycyclic core	Removes hydrogen bond
F439A	Polycyclic core	Removes stacking interactions
S440W	Polycyclic core	Increases steric hindrance
Mut1	*tert*-Butyl ester	Reduces hydrophobic interactions
Mut2	Isobutyl group	Reduces hydrophobic interactions

3.3 Construct generation

Mutant ABCG2 constructs were created by performing mutagenic PCR with pcDNA™3.1/Zeo(+)_SNAP_ABCG2 plasmid as the template DNA which is shown Figure 3.4. Key features of this plasmid include an ampicillin resistance gene (AmpR) for bacterial selection; a Zeocin™ resistance gene (BleoR) for mammalian cell selection; and the cytomegalovirus (CMV) promoter for enhanced expression of the protein of interest, SNAP-tagged ABCG2.

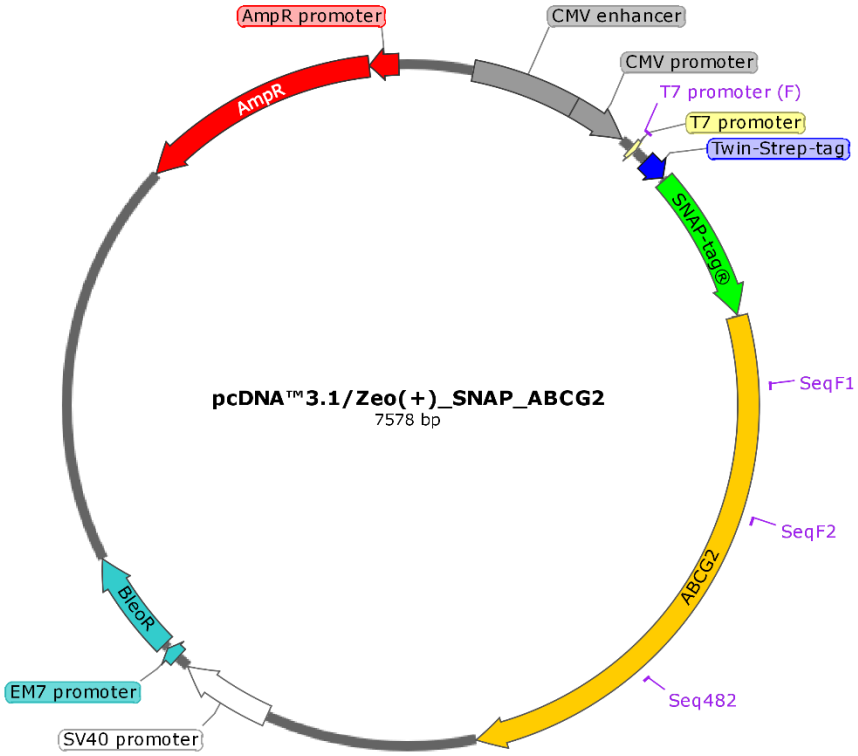


Figure 3.4 Schematic of the template construct used for mutagenic PCR. The pcDNA™3.1/Zeo(+) backbone contains the coding region for twin-strep and SNAP-tagged ABCG2. Key features are shown including AmpR and BleoR to confer ampicillin and Zeocin™ resistance respectively and a CMV promoter for enhanced expression. Primers used in sequencing are shown in purple.

In order to make the mutant constructs, mutagenic PCR was performed using primers containing the mutant codon. The primers were designed to optimise the following factors: ΔG of self-dimer formation greater than -10 kcal/mol; ΔG of hairpin formation greater than -5 kcal/mol; GC content between 40% and 60%; and a score higher than 75% (all found using Premier Biosoft's Netprimer). For the triple mutants Mut1 and Mut2, two sets of primers were designed. The first set made two mutations relatively close in sequence to each other (A397S and V401A for Mut1, I543A and V546A for Mut2). The second set included the final mutation (L539A for Mut1, L405A for Mut2).

PCRs were performed as described in 2.2.1 and resulting DNA was confirmed on agarose gel electrophoresis, shown in Figure 3.5. For the single mutants, S440W, M549E, N436A and T435A have clear bands between 4.0 kb and 10.0 kb which correlates with the construct being 7.6 kb. There is no band, however, for F439A so a gradient PCR was conducted at a range of annealing temperatures (4 temperatures 50.5-65.0 °C), all of which were successful (Figure 3.5 C). Two rounds of PCR were performed for the triple mutants, Mut1 and Mut2. First, the double mutant primers were used (A397S/V401A and I543A/V546A, Table 2.1) on the WT template shown in Figure 3.4. A397S/V401A was successful at the original annealing temperature (55.0 °C, Figure 3.5 B) and I543A/V546A was successful at 50.5 °C (Figure 3.5 C). Once successful PCR was confirmed (Figure 3.5 B and C) and sequences were checked by Sanger sequencing (see below), the PCR products were used as the template in the second round of PCR with the second set of primers (L539A for Mut1 and L405A for Mut2). The resulting PCR products were confirmed by agarose gel electrophoresis shown in Figure 3.5 D.

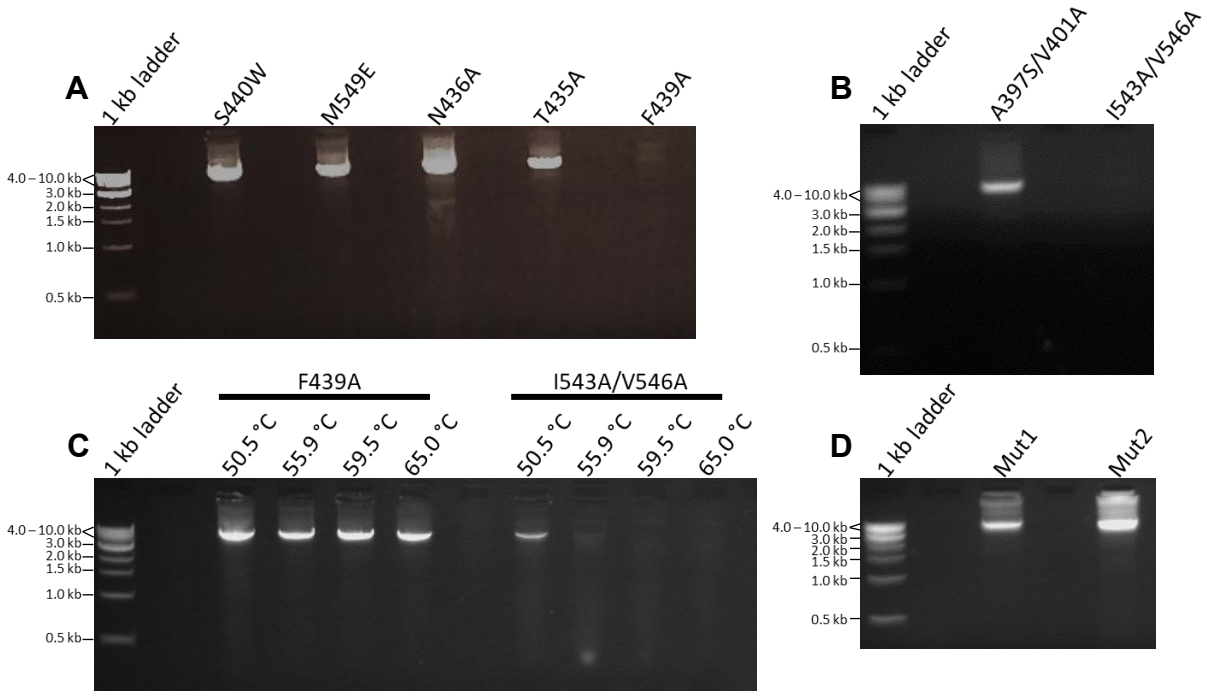


Figure 3.5 Confirmation of successful PCR by gel electrophoresis. PCR was performed using the pcDNA™3.1/Zeo(+)_SNAP_ABCG2 plasmid as the template DNA and the mutagenic primers shown in Table 2.1. **(A)** For the single mutants, all but F439A have undergone successful PCR **(B)** First round of PCR for Mut1 and Mut2, using the double mutant primers to create A397S/V401A and I543A/V546A respectively. A397S/V401A was successful but I543A/V546A was not. **(C)** PCR was performed at a range of annealing temperatures (50.5-65.0 °C) for F439A and I543A/V546A. F439A was successful at all temperatures and I543A/V546A at 50.5 °C **(D)** Second round of PCR for Mut1 and Mut2, using the A397S/V401A and the I543A/V546A (50.5 °C) PCR products respectively as the templates, was successful.

The PCR products were then treated with *DpnI* to break down the methylated template DNA (WT), leaving only the mutated DNA intact. The PCR products were then transformed into DH5α competent *E. coli* and following plasmid isolation, the DNA was sequenced by Source BioScience (Nottingham). Sanger sequencing was performed using a series of primers that covered the entire region of interest (SNAP-labelled ABCG2). Primers used, as well as the region they cover, are shown in Table 2.3 and the position they bind on the plasmid is shown in Figure 3.4. Sequence chromatograms were first viewed in Chromas to confirm quality of the data and then

sequences were compared with the WT construct by aligning in BLAST to see if the mutation was successfully incorporated (Appendix, section 6.2). Successful mutagenesis can be observed for N436A (Figure 3.6, top panel) and similar data for all other single and triple experimental mutants was obtained (Appendix, section 6.1). For N436A there was one unexpected difference between WT and observed sequence (Figure 3.6, bottom panel): an adenine to guanine which changed the TCA codon to TCG. Both these codons code for serine so the construct was still able to be used. Strictly speaking this mutant is N436A/S622S but will be referred to as N436A throughout this thesis. All the other mutant constructs were exactly as expected.

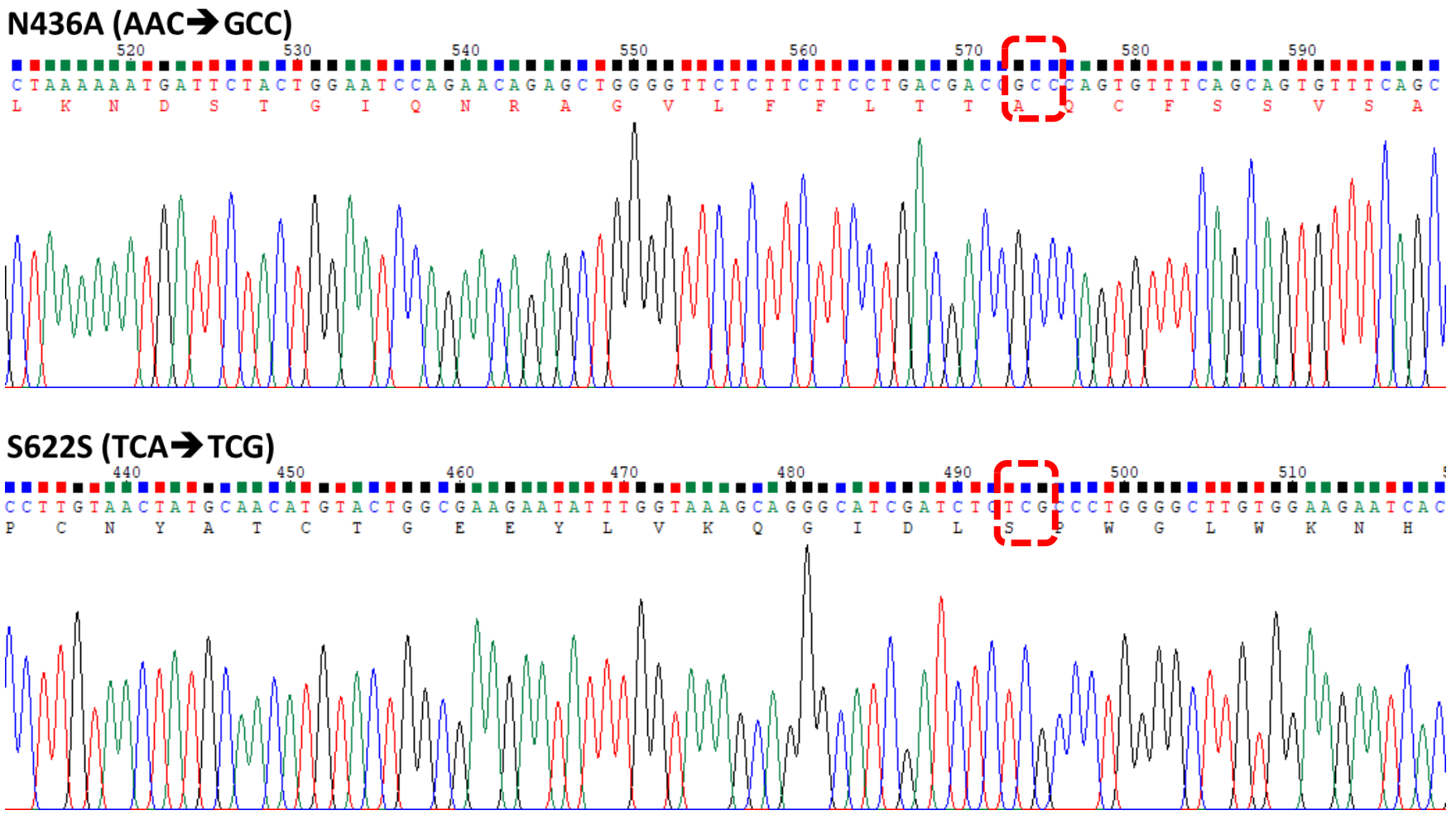


Figure 3.6 Sequencing chromatograms to confirm the desired mutation within the mutant ABCG2 constructs. The top panel shows a portion of the N436A sequencing data obtained using the SeqF2 primer, with the desired mutation highlighted by the red box. The bottom panel shows the additional mutation (red box) in the N436A construct (data collected using the Seq482 primer). The other mutant constructs showed similar results to the top panel, with no additional mutations.

3.4 Protein expression

The mammalian expression system used in this study was HEK293T cells. They are derived from HEK293 cells which were created from normal human embryonic kidney (HEK) cells exposed to sheared fragments of human adenovirus type 5 DNA, making them more infectious. They also contain the temperature-sensitive SV40 T-antigen tsA1609 which activates replication of vectors with the SV40 origin of replication (Figure 3.4). This allows high levels of expression of the protein of interest, SNAP-tagged ABCG2 (Graham et al., 1977, DuBridge et al., 1987, Rio et al., 1985).

HEK293T cells, which have no endogenous ABCG2 expression, were transiently transfected with WT and mutant pcDNA™3.1/Zeo(+)_SNAP_ABCG2 plasmids using the PEI transfection method described by Boussif et al. (1995). The molar PEI nitrogen: DNA phosphorus ratio was 15:1 (section 2.3.2). Stable cell lines were generated by Zeocin™-mediated selection. Zeocin™ concentration in the media was maintained at 200 µg/mL for 36-43 days before reducing to 40 µg/mL. This kills the cells that have not taken up the plasmid, while successfully transfected cells are resistant due to the newly acquired Zeocin™-resistance gene, BleoR (Figure 3.4). The SNAP-labelled ABCG2 is not particularly conducive to rapid observations of expression in live cells; therefore, as a guide for zeocin selection efficacy, the pcDNA™3.1/Zeo(+)_sfGFP_ABCG2 plasmid, coding for GFP-tagged WT-ABCG2, was transfected alongside the WT and mutant plasmids described above. This allowed the expression of ABCG2 to be monitored by epifluorescence microscopy. Figure 3.7 shows that, even after 5 days of Zeocin™ selection, most of the GFP-ABCG2 cells are fluorescent meaning that they are expressing GFP-ABCG2. Observation of untransfected HEK293T cells confirmed that cell death was not

complete until several weeks of exposure. Once stable cell lines were generated, aliquots were frozen in liquid nitrogen for long term storage.

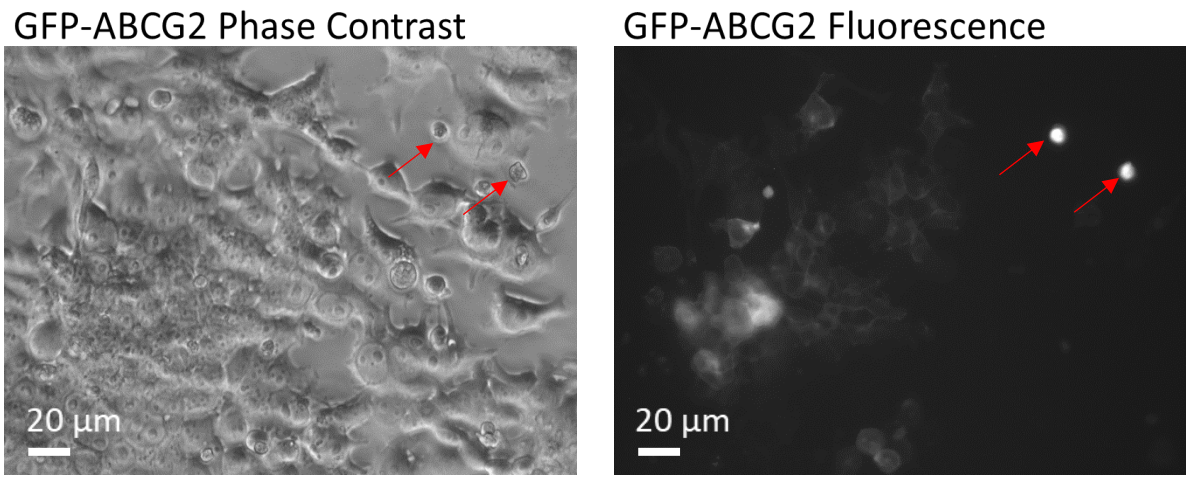


Figure 3.7 Confirmation of GFP fluorescence as an indicator of transfection efficacy. pcDNA™3.1/Zeo(+)_sfGFP_ABCG2 plasmid was transfected into HEK293T cells alongside the WT and mutant constructs described in section 3.2. After 5 days of Zeocin™-selection, epifluorescence microscopy was used to view the GFP-ABCG2 fluorescence of the cells. Fluorescence of these cells indicates that transfection was successful. Cells that are rounded and intensely fluorescent (red arrowheads) are dying/dead cells, presumably reflecting cells that have not taken up the plasmid and so now suffering the cytotoxic effects of zeocin. Scale bars represent approximately 20 µm.

To determine expression levels of ABCG2 in the WT and mutant cell lines, western blotting was performed (Figure 3.8). First cells were harvested, then lysed in ice cold buffer by sonication. 35 µg of cell lysate was resolved electrophoretically on 10% polyacrylamide gel and transferred to nitrocellulose or stained with InstantBlue™. Blots were then probed against the anti-ABCG2 antibody, BXP-21. Figure 3.8 A shows bands of varying intensity between 98 kDa and 148 kDa, which is a larger apparent molecular weight than the expected size of ABCG2 (72.3 kDa) tagged with SNAP-tag® (19.4 kDa) and twin strep tag (3.4 kDa). However, many membrane proteins have mobility differences on SDS-PAGE due to increased interactions with SDS molecules, increasing their apparent molecular weight compared with the

globular protein standards (Rath et al., 2009). F439A and M549E have weak BXP-21 reactive bands alongside lower protein loading demonstrated by the gels stained with InstantBlue™. Increasing the protein loading to 70 µg (Figure 3.8 B) confirmed that F439A was expressed but showed that M549E still had minimal expression of ABCG2. As a result, this mutant was discounted from further experiments. Very low/absent expression of M549E might result from the extreme nature of this mutation, with a hydrophobic methionine being replaced by a hydrophilic glutamic acid, since M549A was successfully expressed in other work (Haider et al., 2015, Manolaridis et al., 2018).

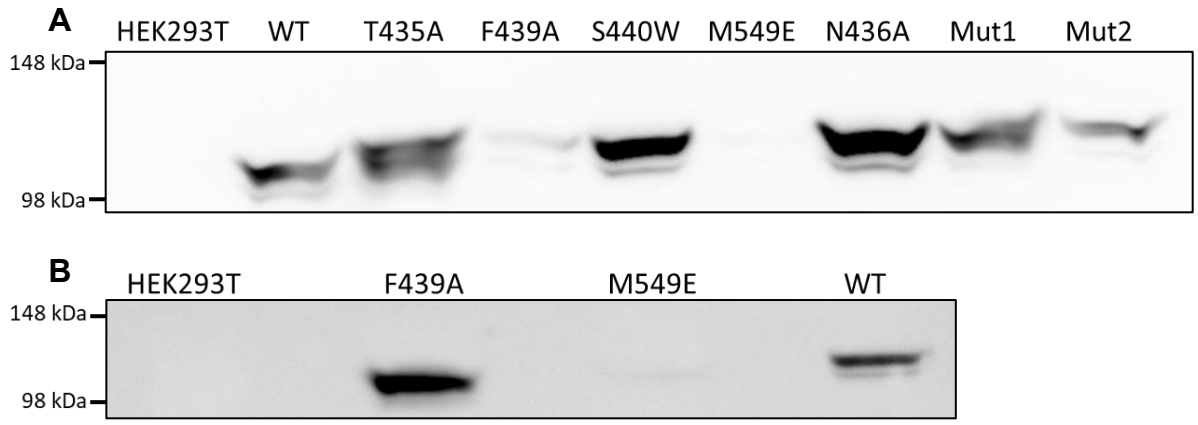


Figure 3.8 Western blots showing mutant and WT-ABCG2 expression in stable cell lines. 35 µg (A) or 70 µg (B) of whole cell lysates underwent SDS-PAGE and immunoblotted against the anti-ABCG2 antibody, BXP-21. Untransfected HEK293T cells were used as a negative control. Change in relative expression of F439A compared to WT can be explained by lower protein loading of F439A and M549E in A and lower protein loading of WT in B. The molecular weights of pre-stained standard proteins were identified from the nitrocellulose.

3.5 Confirmation of protein localisation by flow cytometry

Western blots can confirm total expression of a protein, however, exclusively using this technique for membrane proteins does not provide any information on localisation with the cell. Activity of membrane proteins is influenced by its lipid bilayer surroundings (Lee, 2004, Szilagyi et al., 2017) so localisation on the cell membrane must be confirmed. This is especially true for transporters such as ABCG2, where substrates can only be exported if the protein is correctly localised. In this thesis, flow cytometry was used to quantify the proportion of cells with significant cell surface expression of ABCG2. The monoclonal antibody 5D3 recognises a cell surface epitope within the extracellular loop 3 of ABCG2 and can therefore be used to measure cell surface localisation of ABCG2 (Ozvegy-Laczka et al., 2008). In this project, cells were incubated with 5D3 antibody or isotype control and then the Alexafluor647 secondary antibody and fluorescence was measured by flow cytometry.

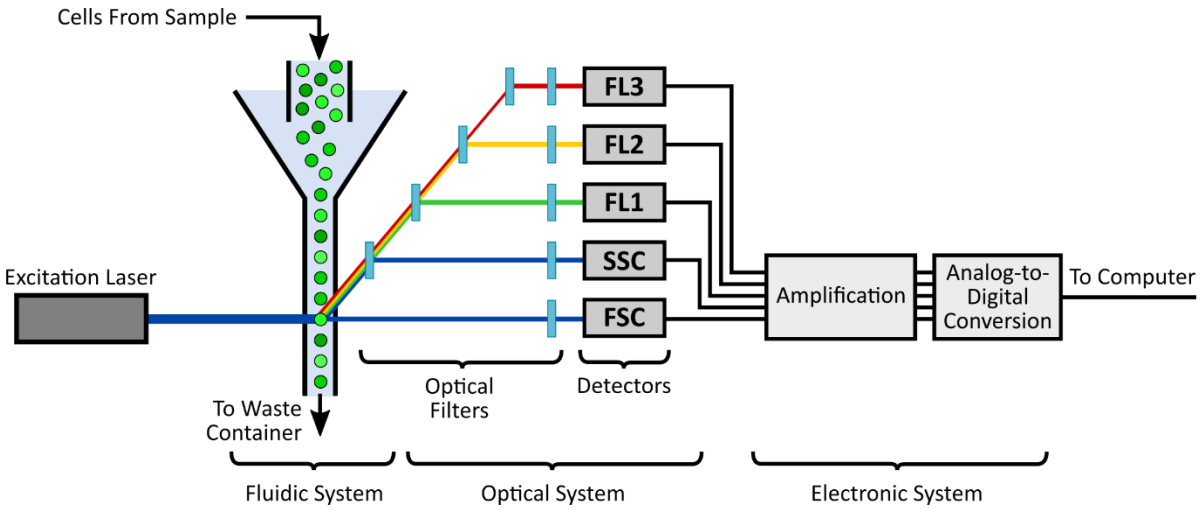


Figure 3.9 Schematic representation of flow cytometry. The basic flow cytometry setup consists of a single stream of cells intercepted by a laser. Light scattered by the cells pass through or are reflected by filters and are detected by side scatter (SSC), forward scatter (SSC) or fluorescent (FL1, FL2, FL3) detectors. Figure acquired from AAT Bioquest (2019).

During flow cytometry, a single stream of cells is produced which allows a laser to illuminate each cell as if flows past (Figure 3.9). The light from the laser is scattered by the cell in all directions but is only detected in the forward direction (same direction as the laser) and at 90° angle from the laser beam (Jaroszeski and Radcliff, 1999). This is referred to as forward scatter (FSC) and side scatter (SSC) respectively. If the cells are fluorescent, either through fluorescent antibodies or proteins (e.g. GFP), the laser will excite the fluorophore and the emitted light will be detected in one of the fluorescent detectors (FL1, FL2, FL3), 90° from the laser. Different wavelengths of emitted light will be either reflected or passed through optical filters so that they are detected by the correct detector.

Side scatter values relates to granularity/structure and forward scatter is proportional to size (Jaroszeski and Radcliff, 1999). This means that different types of cell will appear in different positions when SSC and FSC are plotted against each other and data points can be gated to separate data from each type. This is shown in Figure 3.10 A, where the clear cluster is HEK293T cells and datapoints outside this gate are likely to represent damaged cells and cell debris. Side scatter height plotted against side scatter area can separate monodispersed cells from doublets. Figure 3.10 B shows the gating for singlets, for the untransfected HEK293T cells. This particular sample shows very few doublets but they would appear in a cluster shifted to the right of the one gated. It is important that doublets are gated out because they could give falsely high fluorescence values.

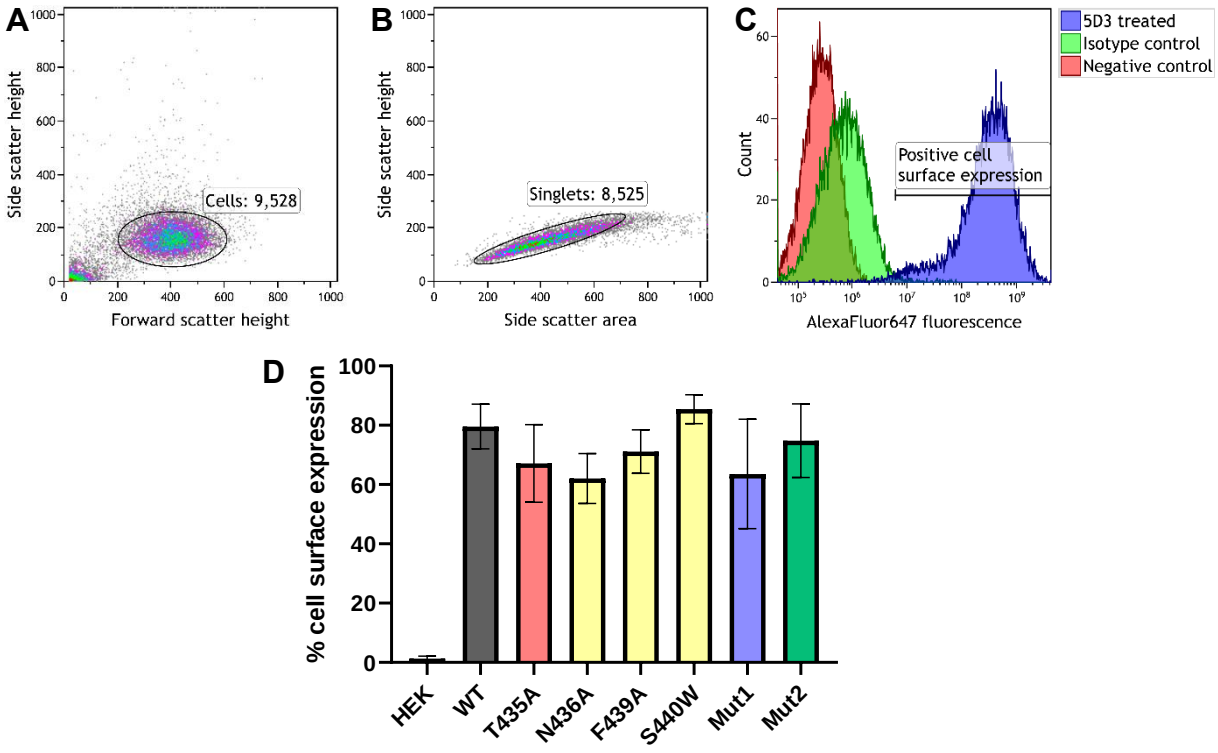


Figure 3.10 Cell surface ABCG2 expression was confirmed by flow cytometry. Cells were treated with either 5D3 (extracellular binding anti-ABCG2 antibody) or isotype control and then AlexaFluor647 secondary antibody. Negative control cells were not treated with any antibody. **(A and B)** Data was first gated to exclude debris from live cells by plotting side scatter height against forward scatter height (A). Side scatter height was then plotted against side scatter area to separate singlets from doublets (B). Number of cells in each gate is shown. Untransfected HEK293T (negative control) is shown and is representative of all data. Density plots are shown, with grey being the sparsest datapoints and red/green being the densest. **(C)** After gating, histograms of cell count vs AlexaFluor647 fluorescence were plotted. C shows the overlay of WT histograms (5D3, isotype control and negative control) for one repeat. It is representative of the remaining data. The right-hand edge of the isotype control peak was used as a threshold for positive cell surface expression for each cell line and repeat. **(D)** Mean percentage of cell surface ABCG2 expression (n=4) of untransfected HEK293T cells (HEK), WT and mutants, derived from 5D3 histograms and isotype control set thresholds. % cell surface expression is the percentage of cells with fluorescence over the threshold set in C. Error bars depict standard error of the mean (SEM). Colours of the bars represent which functional group of Ko143 the mutants interact with (as in Figure 3.3): methoxy (pink), polycyclic core (yellow), *tert*-Butyl ester (blue), isobutyl (green). Control cell lines (HEK and WT) are shown in grey.

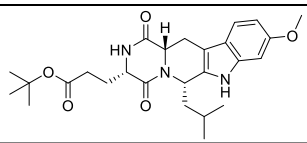
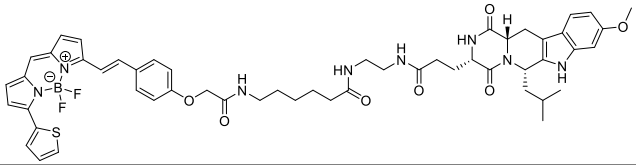
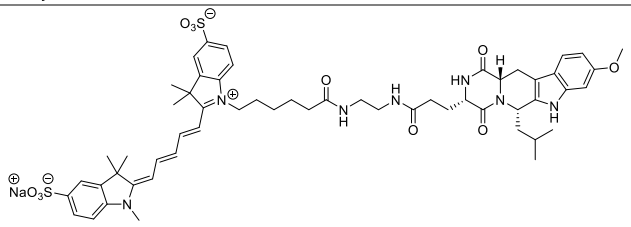
Once data was gated for cells and singlets, histograms for Alexafluor647 fluorescence for each cell line were plotted (Figure 3.10 C). Thresholds for positive cell surface expression were set using the isotype controls for each cell line with any cell displaying greater AlexaFluor647 fluorescence than the 99th centile of the isotype control considered as ABCG2-expressing (Figure 3.10 C). Figure 3.10 D shows the mean percentage of cells with fluorescence over the threshold (% cell surface expression) for all cell lines over 4 repeats. ABCG2 is localised to the plasma membrane to varying degrees but all cell lines (except for untransfected HEK293T cells) have at least 60% cell surface expression and are not significantly different from each other ($p > 0.05$). This is sufficient for use in the transport assays.

3.6 Transport assays

Transport assays have often been performed using flow cytometry and take advantage of many ABCG2 substrates being fluorescent e.g. mitoxantrone, pheophorbide A and Hoechst 33342 (Cox et al., 2018, Kapoor et al., 2020). When cells expressing WT-ABCG2 are incubated with fluorescent substrates, ABCG2 will pump out the substrate leaving a low level of fluorescence in the cell. When treated with inhibitors or using ABCG2 mutants that do not transport the substrate, the substrate will accumulate in the cell giving a higher fluorescence. Ko143, however, is not fluorescent. In order to measure its transport, fluorescent derivatives of Ko143 were used (kindly synthesized by Sarah Mistry, School of Pharmacy, University of Nottingham). Two derivatives, Ko143-X-BY630 and Ko143-Cy5, are shown in Table 3.3 and consist of Ko143 tagged with a fluorophore at the *tert*-Butyl ester end. These will likely bind to ABCG2 in a similar manner to Ko143 with the Ko143-like regions of the fluorescent derivatives binding to cavity 1. All the experimental mutants (except

for Mut1) are designed to reduce affinity for functional groups that are common between Ko143, Ko143-X-BY630 and Ko143-Cy5. Therefore, the effect on binding/transport of the fluorescent derivatives compared with untagged Ko143 should be minimal. However, Mut1 has 3 mutations (A397S/V401A/L539A) all focused around reducing the affinity for the *tert*-Butyl ester of Ko143 (Figure 3.3) but for both fluorescent derivatives the *tert*-Butyl ester is replaced by a fluorescent tag (Table 3.3). Ko143-X-BY630 and Ko143-Cy5 are less hydrophobic than Ko143 in the equivalent region, meaning the experimental mutations would have less of an effect on the affinity than originally intended. In addition, any reduction in affinity caused by a decrease in hydrophobicity, could be cancelled out by the possible formation of a hydrogen bond between A397S and the amides present in the fluorescent Ko143 derivatives.

Table 3.3 Details of Ko143 and its fluorescent derivatives.

Name	Structure	Excitation wavelength (nm)	Emission wavelength (nm)
Ko143		N/A	N/A
Ko143-X-BY630		636	651
Ko143-Cy5		649	666

3.6.1 Monolayer transport assay

Before analysing whether the fluorescent Ko143 derivatives are transported by the mutant ABCG2 proteins, their inhibitory activity compared to Ko143 must be determined. To accomplish this, a monolayer transport assay was performed in collaboration with James Mitchell-White (Kerr Lab, School of Life Sciences, University of Nottingham) using suspension growth-adapted HEK293 (HEK293S) cells which were untransfected or expressing WT-ABCG2 (Lin et al., 2014). Similar to Haider et al. (2015), HEK293S cells were treated with either Hoechst 33342 in the presence or absence of inhibitor (Ko143 or fluorescent derivative). Hoechst 33342 is a fluorescent substrate of ABCG2 whose excitation and emission wavelength (361/497 nm) does not overlap with that of the fluorescent Ko143 derivatives (Table 3.3). The HEK293S and WT cells were also treated with a fluorescent Ko143 derivative (Ko143-X-BY630 or Ko143-Cy5) in the presence or absence of Ko143 (also in absence of Hoechst 33342). The purpose of this is to confirm that the fluorescent derivatives are able to enter HEK293 cells.

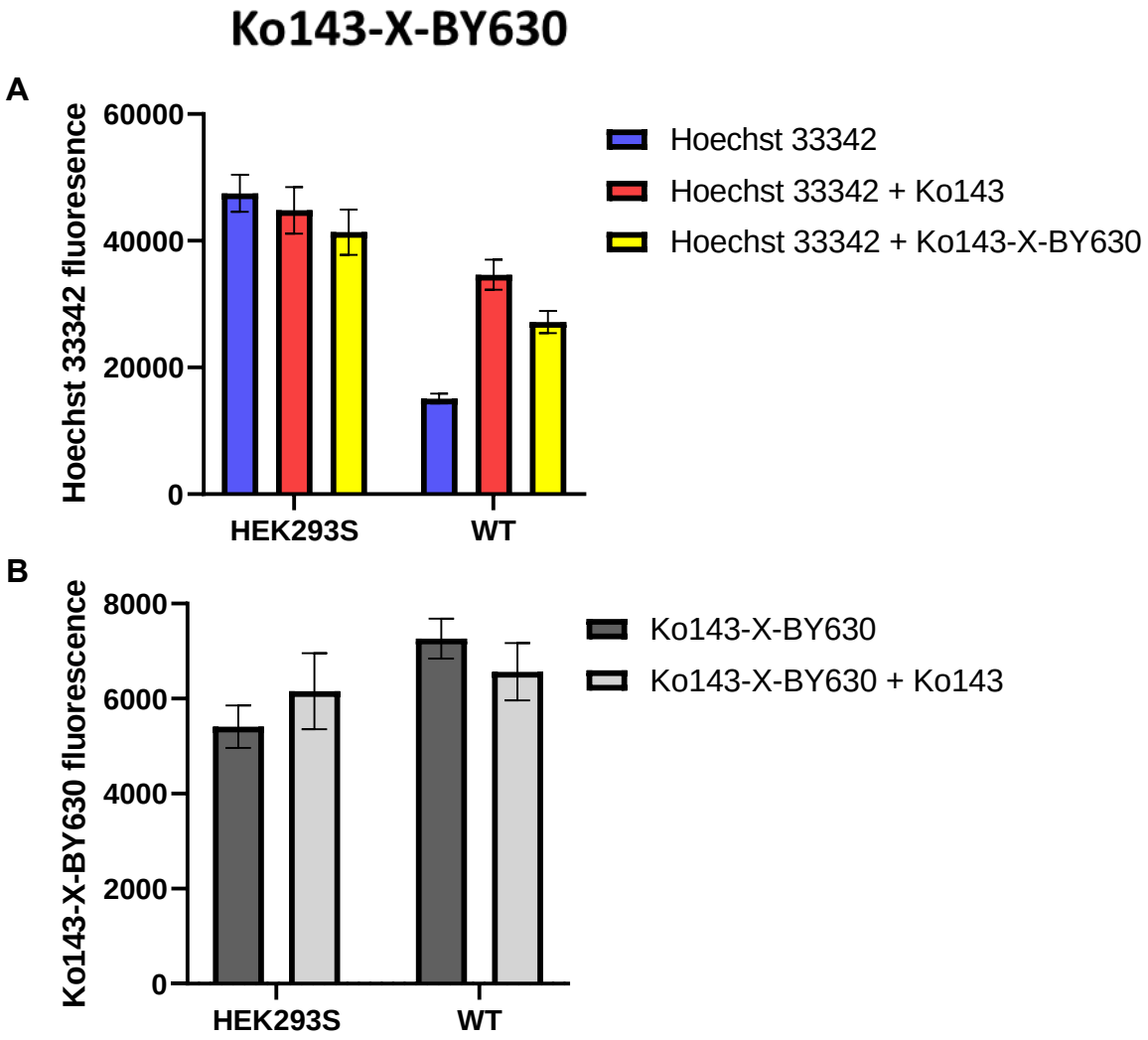


Figure 3.11 Ko143-X-BY630 inhibits Hoechst 33342 transport by ABCG2. (A) Hoechst 33342 fluorescence was measured in untransfected HEK293S and WT expressing HEK293S cells treated with Hoechst 33342 in the presence or absence of an inhibitor (Ko143 or Ko143-BY630). Hoechst 33342 treated WT-ABCG2 expressing cells have significantly less Hoechst 33342 fluorescence than untransfected HEK293S cells. Ko143 and Ko143-X-BY630 inhibit ABCG2 causing a significant increase in Hoechst 33342 accumulation in WT-ABCG2 expressing cells. **(B)** Ko143-X-BY630 fluorescence was measured in untransfected HEK293S and WT expressing HEK293S cells treated with Ko143-X-BY630 in presence or absence of Ko143. All conditions were significantly different from zero showing that Ko143-X-BY630 enters HEK293S and WT-ABCG2 expressing cells. Data also shows no significant Ko143-inhibitable transport of Ko143-BY630. Data was collected on 3 separate occasions. Background fluorescence was accounted for by subtracting the average DMSO fluorescence from each data point (section 2.6). Error bars depict standard error of the mean (SEM).

Figure 3.11 shows the results of the monolayer transport assay for Ko143-X-BY630. When measuring Hoechst 33342 fluorescence (Figure 3.11 A), WT-ABCG2 expressing cells have a significantly different fluorescence ($p < 0.0001$, comparing the two blue bars) to the HEK293S cells because ABCG2 pumps Hoechst 33342 out of the cell. Treatment with Ko143 restores cellular accumulation of Hoechst 33342 giving a higher fluorescence that is not significantly different from the equivalent HEK293S cells. This is because Ko143 inhibits ABCG2 so Hoechst 33342 was not exported. Figure 3.11 A shows that Ko143-X-BY630 acts as inhibitor to almost the same extent as untagged Ko143. Hoechst 33342 transport by WT-ABCG2 expressing cells treated with Ko143-X-BY630 was not significantly different from those treated with Ko143 (compare red and yellow bars, right hand side, Figure 3.11 A) but the magnitude of inhibition appears slightly lower with Ko143-X-BY630. As seen in Figure 3.11 B, Ko143-X-BY630 successfully enters the cell since the Ko143-X-BY630 fluorescence in HEK293S and WT cells is significantly different from zero ($p < 0.0001$). Also, in this case, the addition of the fluorescent tag to Ko143 does not cause it to be transported by WT-ABCG2. This is demonstrated by no significant difference in Ko143-X-BY630 fluorescence between HEK293S and WT cells. In addition, there is no significant difference between Ko143-BY630 treated WT cells in absence or presence of Ko143. This means that, unlike with Hoechst 33342 transport, Ko143 does not affect Ko143-X-BY630 accumulation.

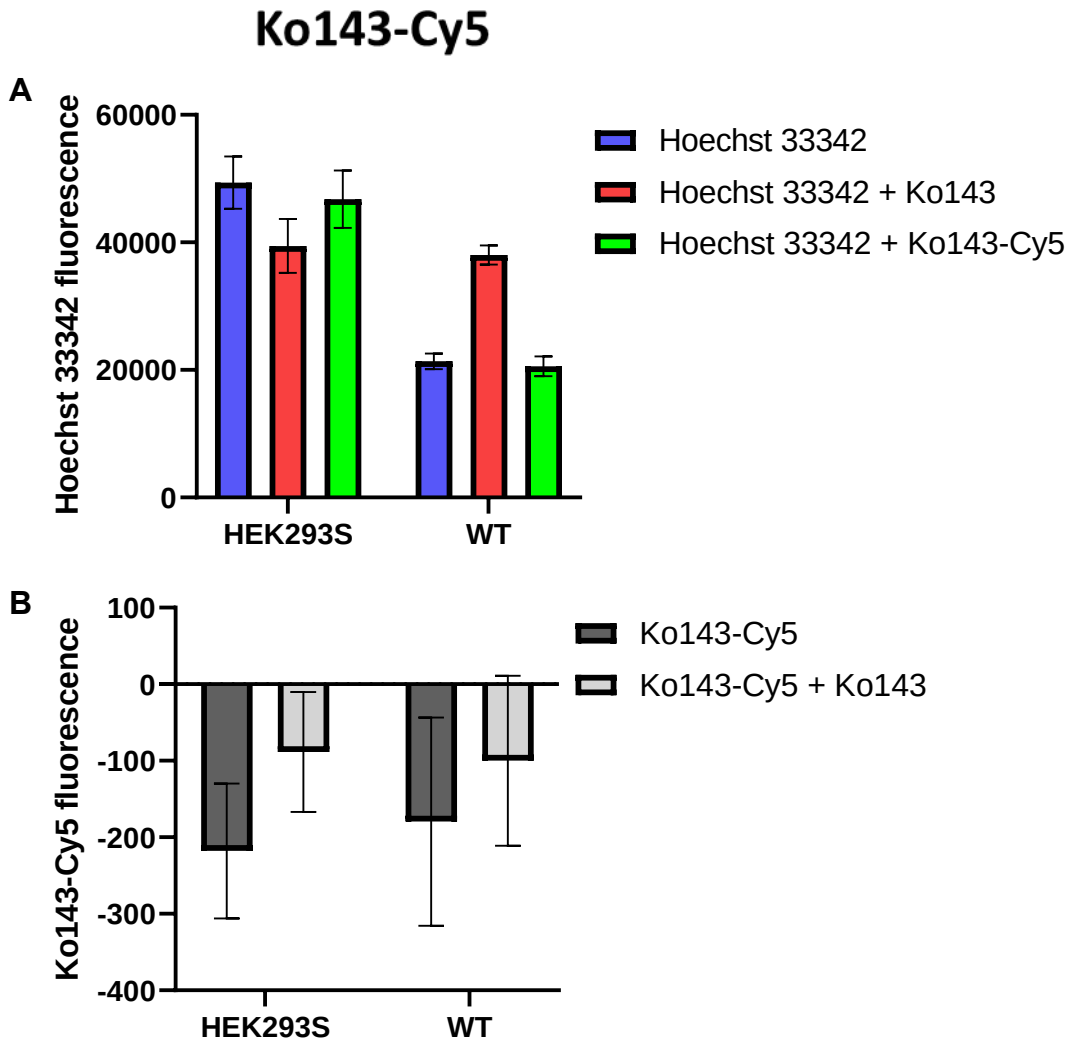


Figure 3.12 Ko143-Cy5 does not inhibit Hoechst 33342 transport by ABCG2. (A) Hoechst 33342 fluorescence was measured in untransfected HEK293S and WT-ABCG2 expressing HEK293S cells treated with Hoechst 33342 in the presence or absence of an inhibitor (Ko143 or Ko143-Cy5). Hoechst 33342 treated WT-ABCG2 expressing cells have significantly less Hoechst 33342 fluorescence than untransfected HEK293S cells. Ko143 inhibits ABCG2 causing a significant increase in Hoechst 33342 accumulation in WT-ABCG2 expressing cells. Ko143-Cy5 does not inhibit Hoechst 33342 transport because fluorescence is not significantly different form WT cells treated with just Hoechst 33342. **(B)** Ko143-Cy5 fluorescence was measured in untransfected HEK293S and WT-ABCG2 expressing HEK293S cells treated with Ko143-Cy5 in presence or absence of Ko143. All conditions were not significantly different from zero showing that Ko143-Cy5 did not enter HEK293S and WT cells. Data was collected on 3 separate occasions. Background fluorescence was accounted for by subtracting the average DMSO from each data point (section 2.6). Error bars depict standard error of the mean (SEM).

The monolayer transport assay was repeated with Ko143-Cy5 instead of Ko143-X-BY630 (Figure 3.12). In contrast to the data for Ko143-X-BY630, Ko143-Cy5 did not act as an inhibitor in the monolayer transport assay. As shown in Figure 3.12 A, there is low cellular accumulation of Hoechst 33342 in WT-ABCG2 expressing cells treated with Ko143-Cy5, which is not significantly different from WT cells treated with Hoechst 33342 alone. This is explained by Figure 3.12 B, where Ko143-Cy5 fluorescence was not significantly different from zero in WT and HEK293S cells. This indicates that Ko143-Cy5 may not successfully enter the cell. Perhaps this is due to the more hydrophilic nature of Ko143-Cy5 compared with Ko143-BY630, having more charged atoms (Table 3.3). This would make it more difficult to cross the largely hydrophobic lipid bilayer of the cell membrane. From this data, it is not possible to determine whether Ko143-Cy5 acts as an inhibitor or not because Ko143-Cy5 did not successfully enter the cell. A membrane-based activity assay, such as the one described in Kapoor et al. (2020) would be able to confirm this ambiguous result.

In summary, the monolayer transport assay measured the cellular accumulation of fluorescent Hoechst 33342 (ABCG2 substrate) and the fluorescent Ko143 derivatives. Ko143-X-BY630 successfully enters the HEK293S cells; inhibits Hoechst 33342 transport by ABCG2 to the same extent as untagged Ko143; and is not transported by ABCG2. Ko143-Cy5, however, did not successfully enter the HEK293S cells and therefore, it is not possible to assess the impact the fluorescent tag has on ABCG2 inhibition.

3.6.2 Flow cytometry

Using the same principle of cellular accumulation as in section 3.6.1, flow cytometry was used to measure whether the mutant cell lines transported the fluorescent Ko143 derivatives. The overarching hypothesis is that if the mutant ABCG2 proteins now recognise Ko143 as a substrate rather than an inhibitor then the accumulation of the fluorescent Ko143 derivatives in those cell lines will be lower. Although the monolayer transport assay indicated that Ko143-Cy5 did not enter cells, the preliminary data suggested that the flow cytometry assay could detect intracellular Ko143-Cy5. Therefore, both fluorescent Ko143 derivatives were used in flow cytometry. The underlying technical differences between the two assays are discussed in Chapter 4.

For the flow cytometry transport assay, mutant cell lines were incubated for 1 hour with 2 μM Ko143-Cy5 or Ko143-X-BY630 and then fluorescence was measured by flow cytometry. WT-ABCG2 expressing cells and untransfected HEK293T (HEK) cells were used as controls as well cells expressing the inactive mutant, E211Q, which has the catalytic glutamic acid in the Walker-B sequence of the NBD mutated to glutamine (Haider et al., 2015), a kind gift of Joseph Morris (Kerr Lab, School of Life Sciences, University of Nottingham). To account for autofluorescence/background fluorescence, control cells for each cell line were treated with DMSO for 1 hour. After gating out debris and doublets (as described in section 3.5), the median fluorescence for each DMSO sample was subtracted from the median fluorescence of the fluorescent Ko143 derivatives. This gave a value for change in median fluorescence intensity (ΔMFI) which were averaged over at least 3 repeats and shown in Figure 3.13.

For Ko143-Cy5 (Figure 3.13 A and Table 3.4), WT-ABCG2 expressing cells have a significantly higher ($p < 0.0001$) ΔMFI than HEK293T cells. This difference is not due to transport of Ko143-Cy5 because HEK293T does not express functionally detectable levels of ABCG2. However, the higher fluorescence of WT cells could be explained by Ko143-Cy5 binding to cavity 1 of ABCG2 leading to cellular retention of the fluorescent inhibitor. All of the cavity 1 mutant cell lines (described in section 3.2), except for Mut1, show significantly different Ko143-Cy5 fluorescence compared to WT-ABCG2 expressing cells ($p < 0.001$) but the same Ko143-Cy5 fluorescence as HEK293T cells. This is possibly due to reduced binding to ABCG2 caused by the designed/purposeful reduction in affinity for Ko143. With this explanation, Ko143-Cy5 would diffuse out of the mutant ABCG2-expressing cells more easily than with the WT-ABCG2 expressing cells, giving a similar fluorescence to HEK293T cells. Altered expression levels of ABCG2 in the mutant cell lines does not appear to be an explanation as Ko143-Cy5 fluorescence does not correlate with expression level of ABCG2 (Figure 3.8). For example, N436A expression was higher than WT-ABCG2 so cannot explain why ΔMFI of N436A expressing cells is lower than WT-ABCG2 expressing cells. The triple mutant Mut1 has a ΔMFI significantly higher than HEK ($p < 0.001$) but not significantly different from WT. As mentioned in section 3.6, the experimental mutations made in Mut1 are focused around reducing the affinity of the *tert*-Butyl ester of Ko143, which is not present in fluorescent Ko143 derivatives. This explains why Mut1 has similar fluorescence to WT cells and supports the idea that the other cavity 1 mutants have reduced binding.

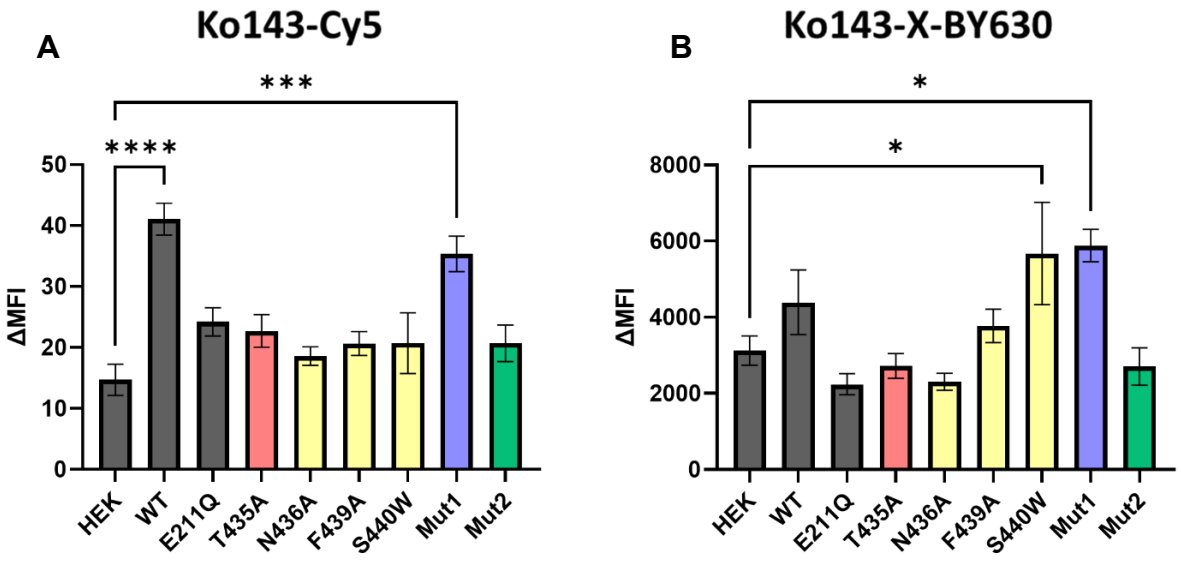


Figure 3.13 Cellular accumulation of fluorescent Ko143 derivatives in mutant cell lines. After gating for monodispersity (see Figure 3.10 A and B), median fluorescence intensity (MFI) was measured for mutant, WT and untransfected cell lines, treated with a fluorescent Ko143 derivative, Ko143-Cy5 (A) or Ko143-X-BY630 (B). Background fluorescence (DMSO control) was subtracted to give ΔMFI. Data was collected on 3 separate occasions for A and 4 separate occasions for B (except F439A where n=3). Colours of the bars represent which functional group of Ko143 the mutants interact with (as in Figure 3.3): methoxy (pink), polycyclic core (yellow), *tert*-Butyl ester (blue), isobutyl (green). Control cell lines (HEK, WT and E211Q) are shown in grey. The lines labelled with * ($p < 0.05$), *** ($p < 0.001$) or **** ($p < 0.0001$) represent mutant cell lines that have a significantly different ΔMFI from HEK293T cells (HEK). Error bars depict standard error of the mean (SEM).

For the Ko143-X-BY630 transport assay (Figure 3.13 B and Table 3.4), the fluorescence level for each cell line shows a similar pattern to the Ko143-Cy5 data. However, S440W is an exception to this and the ΔMFI for all mutant cell lines and HEK293T cells are not significantly different from WT-ABCG2 expressing cells. WT-ABCG2 expressing cells have a ΔMFI that is slightly higher than HEK293T cells along with T435A, N436A, F439A and Mut2 expressing cells having a similar fluorescence to HEK293T cells. As before, Mut1 has a significantly higher ΔMFI than HEK293T cells ($p < 0.05$) but this time the fluorescence is higher (but not significantly) than WT cells. As with Ko143-Cy5, this can be explained by the addition of the fluorescent tag which could the diminish the effect of the hydrophobicity-

reducing experimental mutations. S440W has significantly more Ko143-X-BY630 accumulation than HEK cells ($p < 0.05$, despite the greater variability of this data set), with a fluorescence more similar to Mut1.

Table 3.4 Summary of flow cytometry transport assay results. Significant differences are shown in red (lower) and green (higher) and p values are also shown. Where differences were not significant the phrase "trend to lower/higher" is used.

	Ko143-Cy5		Ko143-X-BY630	
Cell Line	**Compared to WT**	**Compared to HEK**	**Compared to WT**	**Compared to HEK**
HEK293T	Lower ($p < 0.0001$)	N/A	Trend to lower	N/A
WT	N/A	Higher ($p < 0.0001$)	N/A	Trend to higher
E211Q	Lower ($p < 0.01$)	Trend to higher	Trend to lower	Trend to lower
T435A	Lower ($p < 0.001$)	Trend to higher	Trend to lower	Trend to lower
N436A	Lower ($p < 0.0001$)	Trend to higher	Trend to lower	Trend to lower
F439A	Lower ($p < 0.001$)	Trend to higher	Trend to lower	Trend to lower
S440W	Lower ($p < 0.001$)	Trend to higher	Trend to higher	Higher ($p < 0.05$)
Mut1	Trend to lower	Higher ($p < 0.001$)	Trend to higher	Higher ($p < 0.05$)
Mut2	Lower ($p < 0.001$)	Trend to higher	Trend to lower	Trend to higher

In summary, Ko143-Cy5 has significantly higher accumulation in WT-ABCG2 expressing cells compared to HEK293T cells and the other ABCG2 mutants, excluding Mut1. Ko143-X-BY630 shows similar results but without a significant difference (Table 3.4). The exception is S440W which, along with Mut1, has a significantly higher ΔMFI than HEK293T cells. Both Ko143-Cy5 and Ko143-X-BY630 data support the idea that reduced affinity of the mutants for fluorescent Ko143 derivatives causes a decrease in "stickiness" that leads to a lower ΔMFI. Transport of Ko143-Cy5 or Ko143-X-BY630 cannot be ruled out but it is not detectable in this experiment.

Chapter 4 Discussion

At the beginning of this project, a hypothesis was proposed that inhibitors of ABCG2 have a higher affinity for cavity 1 than substrates. It was speculated that reducing the affinity of ABCG2 for Ko143, a well-known inhibitor, would cause it to be transported instead of inhibiting. Data collected was inconclusive with regards to the hypothesis, however, it makes way for additional experiments which could further interrogate the hypothesis. Distinguishing between substrates and inhibitors is important from an academic perspective as well as a clinical one. Not only does it allow better characterisation of substrates which allows drugs to be rationally designed to avoid transport by ABCG2, but also aids development of selective ABCG2 inhibitors. Multidrug resistance is large problem and there are currently no clinically available inhibitors of ABCG2. Known inhibitors such as fumitremorgin C, which is extremely neurotoxic, and Ko143, which is less toxic, are not suitable *in vivo* (Toyoda et al., 2019). If this hypothesis is further validated, ABCG2 inhibitors could be designed by altering non-toxic, selective substrates to have a stronger binding affinity to the transporter. This provides a more guided approach to inhibitor design but also increases the chances of it being clinically suitable since the structurally related substrates have already been approved.

4.1 Summary of results

On a whole, the Ko143-X-BY630 and Ko143-Cy5 flow cytometry data show similar patterns which could be rationalised by the cavity 1 mutants having lower fluorescent Ko143 derivative binding than the WT-ABCG2. However, the flow cytometry data does not rule out the possibility that Ko143-Cy5 or Ko143-X-BY640 are transported by the cavity 1 mutants (except Mut1). If ΔMFI was significantly lower than in

HEK293T cells, this would indicate that the fluorescence Ko143 derivatives were pumped out. Therefore, a reduction in ΔMFI to a value significantly lower than in WT-ABCG2 expressing cells, as observed with most mutants for Ko143-Cy5, could mean that it is transported or that a reduced level of binding to ABCG2 occurs due to the mutations made to cavity 1.

Work by Gose et al. (2020), published after the cavity 1 mutants were designed (section 3.2), confirms a reduction in affinity for Ko143 in ABCG2 mutants F439A and N436A. An aromatic residue 439 was found to be essential for substrate binding and transport, with the authors calling the residue an “aromatic clamp” (Gose et al., 2020). In their work, F439A lost Hoechst 33342 and pheophorbide A transport ability in contrast to their F439W and F439Y mutants. In addition, their thermal stabilisation assay indicated that F439A-ABCG2 reduced binding of multiple substrates and inhibitors including Ko143. Similarly, Manolaridis et al. (2018) found that F439 could be essential for substrate transport because the two F439 residues from both monomers come together which causes cavity 1 to completely collapse. Mutating to a smaller residue could in fact stop substrates being forced out of cavity 1. So, even if it is possible to make Ko143 into a substrate by reducing affinity, this mutant might not be able to transport it regardless. This confirms that in the case of F439A, the most likely scenario is that the fluorescent Ko143 derivatives do not bind ABCG2. This explains why cellular accumulation of Ko143-Cy5 and Ko143-X-BY630 was comparable to untransfected HEK293T cells.

Mutation of N436A by Gose et al. (2020) showed less impact than the F439A mutation in their thermostability assay. This indicated that this residue contributed less to the binding of Ko143, which might be comparable to the lower fluorescence

for Ko143-X-BY630 in the N436A mutant compared to the F439A mutant in the flow cytometry assay (Figure 3.13), although this difference did not reach significance. Therefore, Ko143-X-BY630 could be transported by N436A-ABCG2. Transport by this mutant is possible for some substrates: Hoechst 33342 and pheophorbide A transport activity of N436A is similar to WT but E_1S transport is strongly reduced (Gose et al., 2020, Manolaridis et al., 2018). This supports the idea that difference in fluorescent Ko143 accumulation is due to differences in binding affinity as opposed to increased transport. Gose et al. (2020) suggest that N436A binding selectively to ligands could be due to altered cholesterol modulation of activity or a conformational change which may affect interactions with certain ligands. This could explain why the difference in fluorescence between N436A and F439A is not seen for Ko143-Cy5 but is seen for Ko143-X-BY630.

When treated with Ko143-Cy5, S440W expressing cells showed significantly lower fluorescence than WT-ABCG2 expressing cells and a more comparable fluorescence to untransfected HEK293T cells. In contrast, when treated with Ko143-X-BY630, S440W expressing cells had significantly higher fluorescence than HEK293T cells. This is the exception to the pattern, which on whole matches that of the Ko143-Cy5 flow cytometry data. It is unlikely that there are differences in affinity between the two fluorescent Ko143 derivatives that are not noticed in the other mutants. This is because residue 440 is located in cavity 1 and the structures of Ko143 and both fluorescent Ko143 derivatives would be the same in this location. However, the variability of the data for this particular mutant is high and it is not significantly different from WT-ABCG2 expressing cells, as are the other mutants. The E211Q mutant is not expected to transport anything since the catalytic residue required for ATP hydrolysis has been mutated (Haider et al., 2015, Hollenstein et al., 2007). Its

lower ΔMFI may be partially explained by this mutant consistently expressing at a lower level than WT, meaning fewer ABCG2 protein molecules for Ko143-Cy5 to bind to (Cox, 2019, Kapoor, 2020).

There was a discrepancy between the monolayer transport assay (section 3.6.1) and the flow cytometry transport assay (section 3.6.2) with regards to Ko143-Cy5 fluorescence. The monolayer transport assay had no significant accumulation of Ko143-Cy5 which suggests that Ko143-Cy5 did not enter the cell. This result alone could be rationalised by the altered hydrophobicity of Ko143 by adding the fluorescent tag (Table 3.3). However, Ko143-Cy5 *did* successfully enter the cell in flow cytometry, suggesting cell permeability is merely reduced and not abolished. The assays differ in terms of Ko143-Cy5 concentration (2 μM in flow cytometry vs 1 μM in monolayer transport assay), total duration of Ko143-Cy5 incubation and subsequent washes and second incubations. Additionally, flow cytometry is more sensitive than the monolayer transport assay so it is a better technique for distinguishing low fluorescence levels from the background compared with the microplate reader (Basiji et al., 2007). This could mean that for both techniques, Ko143-Cy5 entered the cells equally well but flow cytometry was just more capable of measuring it. Future work will include performing a Hoechst 33342 transport assay with flow cytometry to assess the inhibitory capability of Ko143-Cy5 and is explained in more detail in section 4.3.

4.2 Alternate ideas and support of the hypothesis

As mentioned in section 3.1, E_1S transport was increased when T435 was mutated to alanine, which suggests that there in an inverse relationship between affinity and maximal transport (Manolaridis et al., 2018). The same mutant was made in this

project but instead of E_1S transport, the potential transport of fluorescent Ko143 derivatives was studied. The accumulation of Ko143-X-BY630 and Ko143-Cy5 in T435A expressing cells in the flow cytometry experiment (Figure 3.13) was lower than in WT-ABCG2 expressing cells and as stated before, this suggests less binding to ABCG2 but does not rule out the possibility of transport. Since T435A has already shown to increase transport of a substrate, it seems plausible that transport of Ko143 is also possible, especially in conjunction with the flow cytometry data. Gose et al. (2020) found that there was a strong correlation between binding affinity of kinase inhibitors for ABCG2 and their IC_{50} for transport inhibition. This supports the idea that there is a spectrum of inhibitors to substrates, where the higher the affinity of a compound, the more inhibitor character it has (Table 3.1). This suggests that perhaps Ko143 can be turned into a substrate and that the lower fluorescence of the cavity 1 mutants (except Mut1) compared with the WT-ABCG2 expressing cells in Figure 3.13 could be due to transport of the fluorescent Ko143 derivatives.

However, there are some other ideas. Kapoor et al. (2018) hypothesised that substrates enter cavity 1 via a surface “access site”, binding to which triggers conformational changes required for ATP binding and hydrolysis. In contrast, they propose that inhibitors bind directly to cavity 1 so ATP hydrolysis does not occur and the inhibitor is not transported. If this is true then reducing the affinity for an inhibitor will only reduce IC_{50} (Gose et al., 2020) and won’t make it a substrate. However, with the hypothesis in this project, ATP hydrolysis would also be inhibited because without the release of the inhibitor from cavity 1, the transport cycle cannot reset. However, phenolic indenoindole inhibitors of ABCG2 stimulate ATPase activity while also inhibiting mitoxantrone transport (Gozzi et al., 2015). Gozzi et al. (2015) suggest that this is due to multiple binding sites but perhaps they enter cavity 1 via the access site

which allows the conformational changes to occur. By combining these ideas, it is possible that some inhibitors can enter via the access site, activating ATPase activity, and others cannot. In relation to the hypothesis in this project, maybe it is only possible for an inhibitor to become a substrate (by decreasing affinity), if it enters via the access site.

Cholesterol has a unique relationship between structure and function when it comes to binding to and modulating ABCG2. At least 20% (w/w) cholesterol in the cell membrane is required for function of ABCG2 (Storch et al., 2007) and Telbisz et al. (2013) found that depletion of cholesterol in Madin Darby canine kidney cells inhibits pheophorbide A transport. In their cryo-EM structures, Jackson et al. (2018) found ordered cholesterol molecules peripheral to ABCG2, which are potentially involved in modulation. Two cholesterol molecules are observed in cavity 1 in the cryo-EM structures by Taylor et al. (2017) and bind in the same location as MZ29, the Ko143 derivative (Jackson et al., 2018, Kerr et al., 2021). The outcome of this binding is unknown since cholesterol is not considered to be a substrate for transport. Perhaps it can act as an inhibitor in cavity 1, preventing the protein from hydrolysing ATP in a so-called “futile cycle” and is also an allosteric modulator at the sites on the protein periphery (Kerr et al., 2021).

New cryo-EM structures of ABCG2 show that part of the transmembrane helix 2 (TM2, residues 434-438) is unwound in the apo-state compared with the substrate/inhibitor-bound state where this region is fully helical (Orlando and Liao, 2020, Jackson et al., 2018, Manolaridis et al., 2018). This unravelling positions some residues important for ligand binding (e.g. F439) away from cavity 1 which is sealed off in this conformation. Since R482, which is situated in the access site and is known

for affecting substrate specificity, interacts with the unravelled portion of TM2, it has been suggested that control over the conformation of TM2 could be a factor in distinguishing between inhibitors and substrates (Orlando and Liao, 2020, Kapoor et al., 2018).

Egido et al. (2015) found that compounds that activate ABCG2 ATPase activity tend to be hydrophilic, non-amphiphilic and highly charged. On the other hand, inhibitors tend to be hydrophobic, amphiphilic and moderately charged to non-charged. In contrast to the hypothesis suggested in this thesis, they propose that the molecular properties relating to amphiphilicity, hydrophobicity, ionisation indicate whether a compound is an inhibitor or substrate (Egido et al., 2015). So, you would be able to predict where the compound falls on the substrate-inhibitor spectrum based on its properties. It would be interesting to examine how these properties correlate to affinity for ABCG2.

4.3 Future experiments

Some experiments are already underway in the Kerr lab with the aim to provide more data to support or disprove the hypothesis. Since flow cytometry had more success with Ko143-Cy5 cell permeability than the monolayer transport assay, a Hoechst 33342 transport assay will be performed using a combined approach of both assays (section 2.5.2 and 2.6). This will include WT and untransfected HEK cells treated with Hoechst 33342 in absence or presence of Ko143 or Ko143-Cy5 as described in Table 2.4 and in Cox et al. (2018) and in Kapoor et al. (2020). If this shows a different result to the monolayer transport assay, a dose response curve will be produced with flow cytometry by measuring inhibition by Ko143, Ko143-Cy5 and Ko143-X-BY630 at different concentrations. This produces a value for half maximal inhibition

concentration (IC_{50}). Comparison of this value for the original Ko143 with the modified derivatives would indicate whether they were more or less potent inhibitor of ABCG2. For the cavity 1 mutants, IC_{50} of untagged Ko143 will be calculated by performing a plate reader accumulation assay as described by Horsey et al. (2020) and measuring mitoxantrone transport at a range of Ko143 concentrations. This would enable quantification of the impact of any one mutant on the interaction of ABCG2 with Ko143.

As for the flow cytometry transport assay (section 2.5.2 and 3.6.2), perhaps adding fluorescent Ko143 in excess will negate the effect of WT having increased binding to fluorescent Ko143. The overall fluorescence of the cell would be higher in non-transporting cell lines so the impact of the “stickiness” of fluorescent Ko143 derivatives to ABCG2 would be decreased. This could potentially decrease the difference between HEK293T cells and the WT-ABCG2 expressing cells and could help clarify whether the decrease in fluorescence from mutant to WT-ABCG2 expressing cells is due to transport of fluorescent Ko143 derivatives or not.

Another method distinguishing between transport and decreased binding would be to prepare inside-out vesicles from cells overexpressing WT and mutant ABCG2 (Karlsson et al., 2010, Toyoda et al., 2019). ABCG2 would then pump substrates and potentially the fluorescent Ko143 derivatives into the vesicle and the fluorescence of these vesicles could be measured to find the accumulation of the fluorescent Ko143 derivatives inside the vesicle. The benefit of this approach is that the cell permeability of Ko143-Cy5 would not be an issue. Ko143-Cy5 in the media would be able interact with cavity 1 of ABCG2 without having to cross the cell membrane first. Also, the medium can be treated with reagents for ATP regeneration or with a non-

hydrolysable ATP analogue AMP-PNP, which allows the measurement of ATP-dependent transport (Karlsson et al., 2010, Toyoda et al., 2019). The difference between vesicle fluorescence in the presence of ATP or AMP-PNP would show how much the fluorescent Ko143 derivatives are transported.

In an ideal world, affinity measurements of the mutant ABCG2 proteins for Ko143 could be compared to WT and a range of substrates, perhaps by measuring thermal shift of ligand binding as described by Gose et al. (2020). This is because there will be a minimum affinity threshold at which transport is no longer possible. It could be possible that the mutations made in this project are too strong and resultant affinity of the Ko143 derivatives is too low to bind, let alone be transported. So even if it is possible to make a substrate out of an inhibitor, we do not know if these experimental mutations are too extreme.

Ko143-X-BY630 inhibits Hoechst 33342 transport to almost the same extent as unmodified Ko143 which means in terms of affinity Ko143-X-BY630 is a good model for how Ko143 would behave in the same circumstances. Yet, the fluorescent tag might still have an impact on transport, potentially being too large or even restricting the conformational changes required for transport. Ko143-Cy5 behaves similarly to Ko143-X-BY630 in the flow cytometry but even after the inhibitory capability is confirmed the tag could prevent transport. Perhaps some inhibitors can never be substrates for the same reason and the hypothesis in this project might only apply to some inhibitors. Radiolabelling Ko143 with a radioactive atom, such as ^{3}H, would be a suitable method to overcome the effect of the large fluorescent tags.

If the hypothesis in this project is ever supported by more conclusive data, the next step would be to make chemical modifications to existing substrates to increase

affinity for ABCG2. Not only does this have the potential for creating non-toxic inhibitors from substrates such as E_1S, but perhaps altering drugs that are affected by multidrug resistance (e.g. mitoxantrone or methotrexate, (Doyle and Ross, 2003)) might also prevent their own transport.

4.4 Conclusion

Without further experiments there is currently ambiguity about whether the mutants made lead to Ko143 transport. However, it is likely that lower fluorescence of the cavity 1 mutants is due to reduced binding to the fluorescent Ko143 derivatives or perhaps a small level of transport. Data from Gose et al. (2020) compliments the hypothesis in this thesis, showing more potent inhibitors have a higher affinity than weaker inhibitors. Mutations to introduce a smaller reduction in affinity or measuring the affinity for Ko143 in the current mutants will help clarify if the experimental mutations are too strong to transport. It is potentially possible for inhibitors to become substrates but other factors, such as access site binding, could also influence substrate/inhibitor quality which prevents some inhibitors being substrates.

Chapter 5 References

AAT BIOQUEST. 2019. *Fundamentals of Flow Cytometry* [Online]. Available: https://www.aatbio.com/resources/assaywise/2019-8-1/fundamentals-of-flow-cytometry [Accessed 30/04/2021].

ALLEN, J. D., VAN LOEVEZIJN, A., LAKHAI, J. M., VAN DER VALK, M., VAN TELLINGEN, O., REID, G., SCHELLENS, J. H., KOOMEN, G. J. & SCHINKEL, A. H. 2002. Potent and specific inhibition of the breast cancer resistance protein multidrug transporter in vitro and in mouse intestine by a novel analogue of fumitremorgin C. *Mol Cancer Ther,* 1**,** 417-25.

BASIJI, D. A., ORTYN, W. E., LIANG, L., VENKATACHALAM, V. & MORRISSEY, P. 2007. Cellular image analysis and imaging by flow cytometry. *Clin Lab Med,* 27**,** 653-70, viii.

BOUSSIF, O., LEZOUALC'H, F., ZANTA, M. A., MERGNY, M. D., SCHERMAN, D., DEMENEIX, B. & BEHR, J. P. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A,* 92**,** 7297-301.

CHIANG, J. Y. 2003. Bile acid regulation of hepatic physiology: III. Bile acids and nuclear receptors. *Am J Physiol Gastrointest Liver Physiol,* 284**,** G349-56.

COX, M. H. 2019. *ABCG2: the lateral slice hypothesis as a model for multidrug transport.* PhD Thesis, University of Nottingham.

COX, M. H., KAPOOR, P., BRIGGS, D. A. & KERR, I. D. 2018. Residues contributing to drug transport by ABCG2 are localised to multiple drug-binding pockets. *Biochem J,* 475**,** 1553-1567.

DE LERA RUIZ, M. & KRAUS, R. L. 2015. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. *J Med Chem,* 58**,** 7093-118.

DIOP, N. K. & HRYCYNA, C. A. 2005. N-Linked glycosylation of the human ABC transporter ABCG2 on asparagine 596 is not essential for expression, transport activity, or trafficking to the plasma membrane. *Biochemistry,* 44**,** 5420-9.

DOYLE, L. & ROSS, D. D. 2003. Multidrug resistance mediated by the breast cancer resistance protein BCRP (ABCG2). *Oncogene,* 22**,** 7340-58.

DUBRIDGE, R. B., TANG, P., HSIA, H. C., LEONG, P. M., MILLER, J. H. & CALOS, M. P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular and Cellular Biology,* 7**,** 379-387.

EGIDO, E., MÜLLER, R., LI-BLATTER, X., MERINO, G. & SEELIG, A. 2015. Predicting Activators and Inhibitors of the Breast Cancer Resistance Protein (ABCG2) and P-Glycoprotein (ABCB1) Based on Mechanistic Considerations. *Mol Pharm,* 12**,** 4026-37.

GOSE, T., SHAFI, T., FUKUDA, Y., DAS, S., WANG, Y., ALLCOCK, A., GAVAN MCHARG, A., LYNCH, J., CHEN, T., TAMAI, I., SHELAT, A., FORD, R. C. & SCHUETZ, J. D. 2020. ABCG2 requires a single aromatic amino acid to "clamp" substrates and inhibitors into the binding pocket. *Faseb j,* 34**,** 4890-4903.

GOZZI, G. J., BOUAZIZ, Z., WINTER, E., DAFLON-YUNES, N., HONORAT, M., GURAGOSSIAN, N., MARMINON, C., VALDAMERI, G., BOLLACKE, A., GUILLON, J., PINAUD, N., MARCHIVIE, M., CADENA, S. M., JOSE, J., LE BORGNE, M. & DI PIETRO, A. 2015. Phenolic indeno[1,2-b]indoles as ABCG2-selective potent and non-toxic inhibitors stimulating basal ATPase activity. *Drug Des Devel Ther,* 9**,** 3481-95.

GRAHAM, F. L., SMILEY, J., RUSSELL, W. C. & NAIRN, R. 1977. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology,* 36**,** 59-72.

HAIDER, A. J., COX, M. H., JONES, N., GOODE, A. J., BRIDGE, K. S., WONG, K., BRIGGS, D. & KERR, I. D. 2015. Identification of residues in ABCG2 affecting protein trafficking and drug transport, using co-evolutionary analysis of ABCG sequences. *Biosci Rep,* 35.

HALWACHS, S., KNEUER, C., GOHLSCH, K., MULLER, M., RITZ, V. & HONSCHA, W. 2016. The ABCG2 efflux transporter from rabbit placenta: Cloning and functional characterization. *Placenta,* 38**,** 8-15.

HEGYI, Z. & HOMOLYA, L. 2016. Functional Cooperativity between ABCG4 and ABCG1 Isoforms. *PLoS One,* 11**,** e0156516.

HENRIKSEN, U., FOG, J. U., LITMAN, T. & GETHER, U. 2005. Identification of intra- and intermolecular disulfide bridges in the multidrug resistance transporter ABCG2. *J Biol Chem,* 280**,** 36926-34.

HÖGEL, P., GÖTZ, A., KUHNE, F., EBERT, M., STELZER, W., RAND, K. D., SCHARNAGL, C. & LANGOSCH, D. 2018. Glycine Perturbs Local and Global Conformational Flexibility of a Transmembrane Helix. *Biochemistry,* 57**,** 1326-1337.

HOLLENSTEIN, K., DAWSON, R. J. & LOCHER, K. P. 2007. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol,* 17**,** 412-8.

HOMOLYA, L., ORBÁN, T. I., CSANÁDY, L. & SARKADI, B. 2011. Mitoxantrone is expelled by the ABCG2 multidrug transporter directly from the plasma membrane. *Biochim Biophys Acta,* 1808**,** 154-63.

HORSEY, A. J., BRIGGS, D. A., HOLLIDAY, N. D., BRIDDON, S. J. & KERR, I. D. 2020. Application of fluorescence correlation spectroscopy to study substrate binding in styrene maleic acid lipid copolymer encapsulated ABCG2. *Biochim Biophys Acta Biomembr,* 1862**,** 183218.

HORSEY, A. J., COX, M. H., SARWAT, S. & KERR, I. D. 2016. The multidrug transporter ABCG2: still more questions than answers. *Biochem Soc Trans,* 44**,** 824-30.

JACKSON, S. M., MANOLARIDIS, I., KOWAL, J., ZECHNER, M., TAYLOR, N. M. I., BAUSE, M., BAUER, S., BARTHOLOMAEUS, R., BERNHARDT, G., KOENIG, B., BUSCHAUER, A., STAHLBERG, H., ALTMANN, K. H. & LOCHER, K. P. 2018. Structural basis of small-molecule inhibition of human multidrug transporter ABCG2. *Nat Struct Mol Biol,* 25**,** 333-340.

JAROSZESKI, M. J. & RADCLIFF, G. 1999. Fundamentals of flow cytometry. *Mol Biotechnol,* 11**,** 37-53.

JONES, P. M. & GEORGE, A. M. 1999. Subunit interactions in ABC transporters: towards a functional architecture. *FEMS Microbiol Lett,* 179**,** 187-202.

KAPOOR, P. 2020. *Elucidating the molecular mechanism of drug binding and translocation by the multidrug transporter ABCG2.* PhD Thesis, University of Nottingham.

KAPOOR, P., BRIGGS, D. A., COX, M. H. & KERR, I. D. 2020. Disruption of the Unique ABCG-Family NBD:NBD Interface Impacts Both Drug Transport and ATP Hydrolysis. *Int J Mol Sci,* 21.

KAPOOR, P., HORSEY, A. J., COX, M. H. & KERR, I. D. 2018. ABCG2: does resolving its structure elucidate the mechanism? *Biochem Soc Trans,* 46**,** 1485-1494.

KARLSSON, J. E., HEDDLE, C., ROZKOV, A., ROTTICCI-MULDER, J., TUVESSON, O., HILGENDORF, C. & ANDERSSON, T. B. 2010. High-activity p-glycoprotein, multidrug resistance protein 2, and breast cancer resistance protein membrane vesicles prepared from transiently transfected human embryonic kidney 293-epstein-barr virus nuclear antigen cells. *Drug Metab Dispos,* 38**,** 705-14.

KERR, I. D., HAIDER, A. J. & GELISSEN, I. C. 2011. The ABCG family of membrane-associated transporters: you don't have to be big to be mighty. *Br J Pharmacol,* 164**,** 1767-79.

KERR, I. D., HUTCHISON, E., GERARD, L., ALEIDI, S. M. & GELISSEN, I. C. 2021. Mammalian ABCG-transporters, sterols and lipids: To bind perchance to transport? *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1866**,** 158860.

KIM, Y. & CHEN, J. 2018. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science,* 359**,** 915-919.

KOHLER, S. C. & WIESE, M. 2015. HM30181 Derivatives as Novel Potent and Selective Inhibitors of the Breast Cancer Resistance Protein (BCRP/ABCG2). *J Med Chem,* 58**,** 3910-21.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.

LÁSZLÓ, L., SARKADI, B. & HEGEDŰS, T. 2016. Jump into a New Fold-A Homology Based Model for the ABCG2/BCRP Multidrug Transporter. *PLoS One,* 11**,** e0164426.

LEE, A. G. 2004. How lipids affect the activities of integral membrane proteins. *Biochim Biophys Acta,* 1666**,** 62-87.

LIN, Y. C., BOONE, M., MEURIS, L., LEMMENS, I., VAN ROY, N., SOETE, A., REUMERS, J., MOISSE, M., PLAISANCE, S., DRMANAC, R., CHEN, J., SPELEMAN, F., LAMBRECHTS, D., VAN DE PEER, Y., TAVERNIER, J. & CALLEWAERT, N. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Commun,* 5**,** 4767.

MANOLARIDIS, I., JACKSON, S. M., TAYLOR, N. M. I., KOWAL, J., STAHLBERG, H. & LOCHER, K. P. 2018. Cryo-EM structures of a human ABCG2 mutant trapped in ATP-bound and substrate-bound states. *Nature,* 563**,** 426-430.

MANZINI, L., HALWACHS, S., GIROLAMI, F., BADINO, P., HONSCHA, W. & NEBBIA, C. 2017. Interaction of mammary bovine ABCG2 with AFB1 and its metabolites and regulation by PCB 126 in a MDCKII in vitro model. *J Vet Pharmacol Ther,* 40**,** 591-598.

MCDEVITT, C. A., CROWLEY, E., HOBBS, G., STARR, K. J., KERR, I. D. & CALLAGHAN, R. 2008. Is ATP binding responsible for initiating drug translocation by the multidrug transporter ABCG2? *Febs j,* 275**,** 4354-62.

MIYATA, H., TAKADA, T., TOYODA, Y., MATSUO, H., ICHIDA, K. & SUZUKI, H. 2016. Identification of Febuxostat as a New Strong ABCG2 Inhibitor: Potential Applications and Risks in Clinical Situations. *Front Pharmacol,* 7**,** 518.

MO, W. & ZHANG, J. T. 2012. Human ABCG2: structure, function, and its role in multidrug resistance. *Int J Biochem Mol Biol,* 3**,** 1-27.

NAKAGAWA, H., WAKABAYASHI-NAKAO, K., TAMURA, A., TOYODA, Y., KOSHIBA, S. & ISHIKAWA, T. 2009. Disruption of N-linked glycosylation enhances ubiquitin-mediated proteasomal degradation of the human ATP-binding cassette transporter ABCG2. *Febs j,* 276**,** 7237-52.

NAKATOMI, K., YOSHIKAWA, M., OKA, M., IKEGAMI, Y., HAYASAKA, S., SANO, K., SHIOZAWA, K., KAWABATA, S., SODA, H., ISHIKAWA, T., TANABE, S. & KOHNO, S. 2001. Transport of 7-ethyl-10-hydroxycamptothecin (SN-38) by breast cancer resistance protein ABCG2 in human lung cancer cells. *Biochem Biophys Res Commun,* 288**,** 827-32.

OCHOA-PUENTES, C., BAUER, S., KUHNLE, M., BERNHARDT, G., BUSCHAUER, A. & KONIG, B. 2013. Benzanilide-Biphenyl Replacement: A Bioisosteric Approach to Quinoline Carboxamide-Type ABCG2 Modulators. *ACS Med Chem Lett,* 4**,** 393-6.

ORLANDO, B. J. & LIAO, M. 2020. ABCG2 transports anticancer drugs via a closed-to-open switch. *Nat Commun,* 11**,** 2264.

OZVEGY-LACZKA, C., KÖBLÖS, G., SARKADI, B. & VÁRADI, A. 2005. Single amino acid (482) variants of the ABCG2 multidrug transporter: major differences in transport capacity and substrate recognition. *Biochimica et biophysica acta,* 1668**,** 53-63.

OZVEGY-LACZKA, C., LACZKÓ, R., HEGEDUS, C., LITMAN, T., VÁRADY, G., GODA, K., HEGEDUS, T., DOKHOLYAN, N. V., SORRENTINO, B. P., VÁRADI, A. & SARKADI, B. 2008. Interaction with the 5D3 monoclonal antibody is regulated by intramolecular rearrangements but not by covalent dimer formation of the human ABCG2 multidrug transporter. *J Biol Chem,* 283**,** 26059-70.

RATH, A., GLIBOWICKA, M., NADEAU, V. G., CHEN, G. & DEBER, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A,* 106**,** 1760-5.

RIO, D. C., CLARK, S. G. & TJIAN, R. 1985. A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction. *Science,* 227**,** 23-8.

ROBEY, R. W., STEADMAN, K., POLGAR, O. & BATES, S. E. 2005. ABCG2-mediated transport of photosensitizers: potential impact on photodynamic therapy. *Cancer Biol Ther,* 4**,** 187-94.

STORCH, C. H., EHEHALT, R., HAEFELI, W. E. & WEISS, J. 2007. Localization of the human breast cancer resistance protein (BCRP/ABCG2) in lipid rafts/caveolae and modulation of its activity by cholesterol in vitro. *J Pharmacol Exp Ther,* 323**,** 257-64.

SUZUKI, M., SUZUKI, H., SUGIMOTO, Y. & SUGIYAMA, Y. 2003. ABCG2 transports sulfated conjugates of steroids and xenobiotics. *J Biol Chem,* 278**,** 22644-9.

SZILAGYI, J. T., VETRANO, A. M., LASKIN, J. D. & ALEKSUNES, L. M. 2017. Localization of the placental BCRP/ABCG2 transporter to lipid rafts: Role for cholesterol in mediating efflux activity. *Placenta,* 55**,** 29-36.

TAMURA, A., WAKABAYASHI, K., ONISHI, Y., TAKEDA, M., IKEGAMI, Y., SAWADA, S., TSUJI, M., MATSUDA, Y. & ISHIKAWA, T. 2007. Re-evaluation and functional classification of non-synonymous single nucleotide polymorphisms of the human ATP-binding cassette transporter ABCG2. *Cancer Sci,* 98**,** 231-9.

TAYLOR, N. M. I., MANOLARIDIS, I., JACKSON, S. M., KOWAL, J., STAHLBERG, H. & LOCHER, K. P. 2017. Structure of the human multidrug transporter ABCG2. *Nature,* 546**,** 504-509.

TELBISZ, Á., ÖZVEGY-LACZKA, C., HEGEDŰS, T., VÁRADI, A. & SARKADI, B. 2013. Effects of the lipid environment, cholesterol and bile acids on the function of the purified and reconstituted human ABCG2 protein. *Biochem J,* 450**,** 387-95.

TIRAT, A., FREULER, F., STETTLER, T., MAYR, L. M. & LEDER, L. 2006. Evaluation of two novel tag-based labelling technologies for site-specific modification of proteins. *Int J Biol Macromol,* 39**,** 66-76.

TOWBIN, H., STAEHELIN, T. & GORDON, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A,* 76**,** 4350-4.

TOYODA, Y., TAKADA, T. & SUZUKI, H. 2019. Inhibitors of Human ABCG2: From Technical Background to Recent Updates With Clinical Implications. *Front Pharmacol,* 10**,** 208.

VOLK, E. L. & SCHNEIDER, E. 2003. Wild-type breast cancer resistance protein (BCRP/ABCG2) is a methotrexate polyglutamate transporter. *Cancer Res,* 63**,** 5538-43.

WANG, N., YVAN-CHARVET, L., LÜTJOHANN, D., MULDER, M., VANMIERLO, T., KIM, T. W. & TALL, A. R. 2008. ATP-binding cassette transporters G1 and G4 mediate cholesterol and desmosterol efflux to HDL and regulate sterol accumulation in the brain. *Faseb j,* 22**,** 1073-82.

WEI, J., GEALE, P. F., SHEEHY, P. A. & WILLIAMSON, P. 2012. The impact of ABCG2 on bovine mammary epithelial cell proliferation. *Anim Biotechnol,* 23**,** 221-4.

WEIDNER, L. D., ZOGHBI, S. S., LU, S., SHUKLA, S., AMBUDKAR, S. V., PIKE, V. W., MULDER, J., GOTTESMAN, M. M., INNIS, R. B. & HALL, M. D. 2015. The Inhibitor Ko143 Is Not Specific for ABCG2. *J Pharmacol Exp Ther,* 354**,** 384-93.

WONG, K., BRIDDON, S. J., HOLLIDAY, N. D. & KERR, I. D. 2016. Plasma membrane dynamics and tetrameric organisation of ABCG2 transporters in mammalian cells revealed by single particle imaging techniques. *Biochim Biophys Acta,* 1863**,** 19-29.

WOODWARD, O. M., KÖTTGEN, A., CORESH, J., BOERWINKLE, E., GUGGINO, W. B. & KÖTTGEN, M. 2009. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proc Natl Acad Sci U S A,* 106**,** 10338-42.

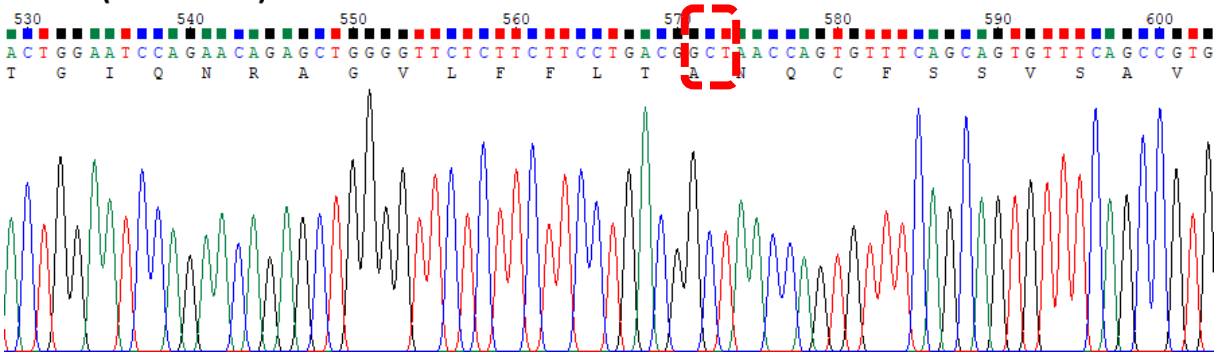
ZHANG, H., CATANIA, R. & JEUKEN, L. J. C. 2020. Membrane Protein Modified Electrodes in Bioelectrocatalysis. *Catalysts,* 10**,** 1427.

Chapter 6 Appendix

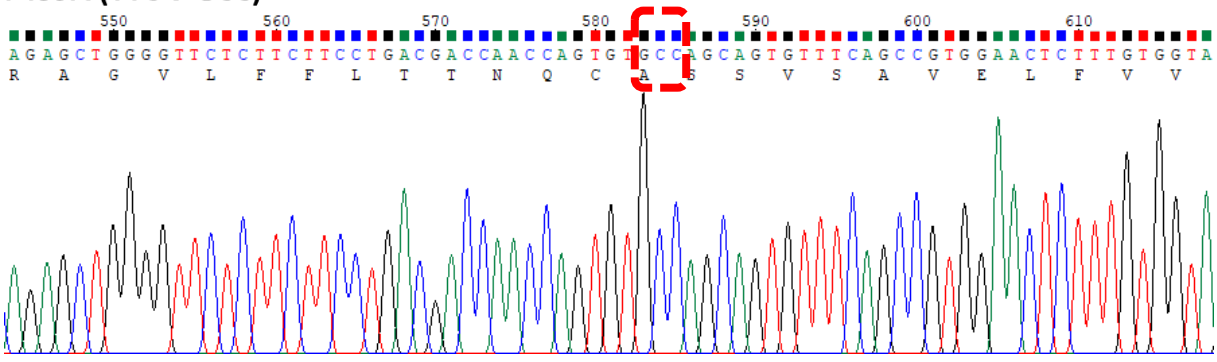
6.1 Sequencing chromatograms

Portions of the sequencing chromatograms of the remaining mutant ABCG2 constructs (N436A is shown in Figure 3.6) are shown below. The SeqF2 primer was used for T435A, F439A, S440W, A397S/V401A (Mut1) and L405A (Mut2). The Seq482 primer was used for M549E, L539A (Mut1) and I543A/V546A (Mut2). The red box highlights the desired mutation and no other mutations were found.

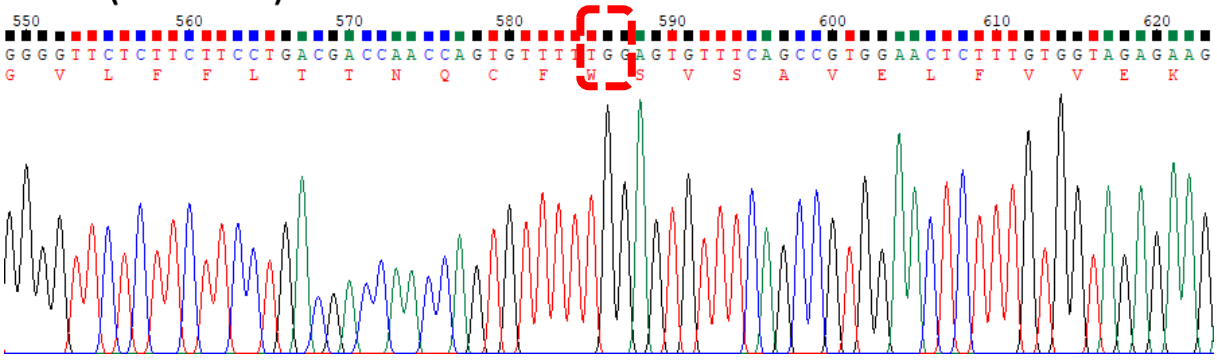
T435A (ACC→GCT)



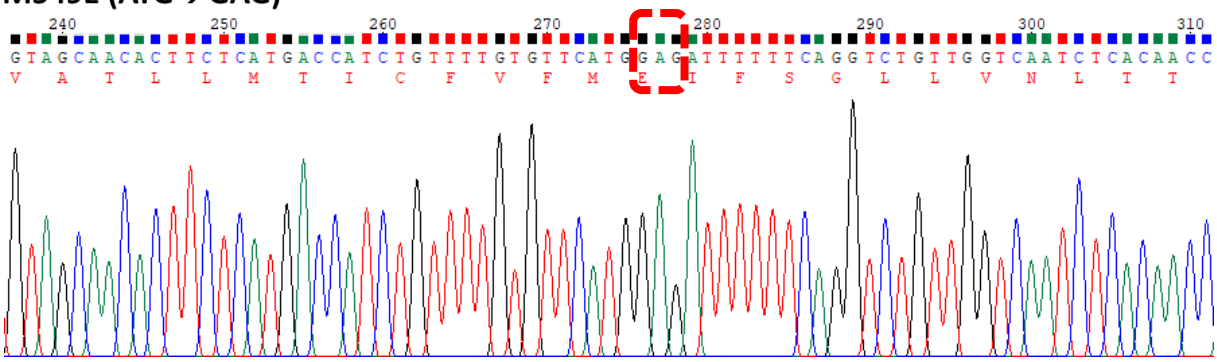
F439A (TTC→GCC)



S440W (AGC→TGG)

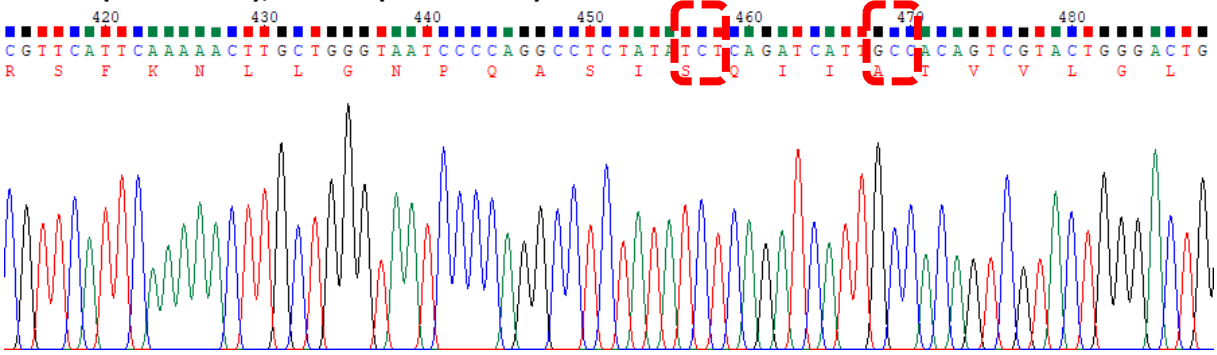


M549E (ATG→GAG)



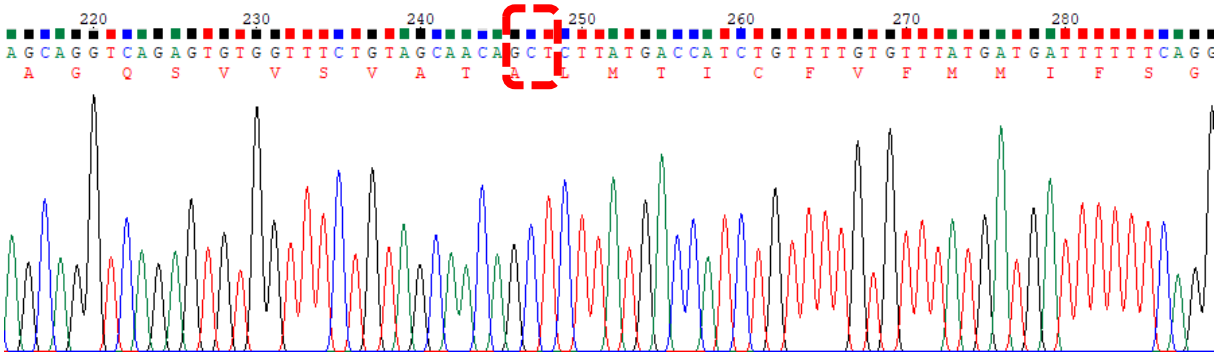
Mut1

A397S (GCT→TCT), V401A (GTC→GCC)



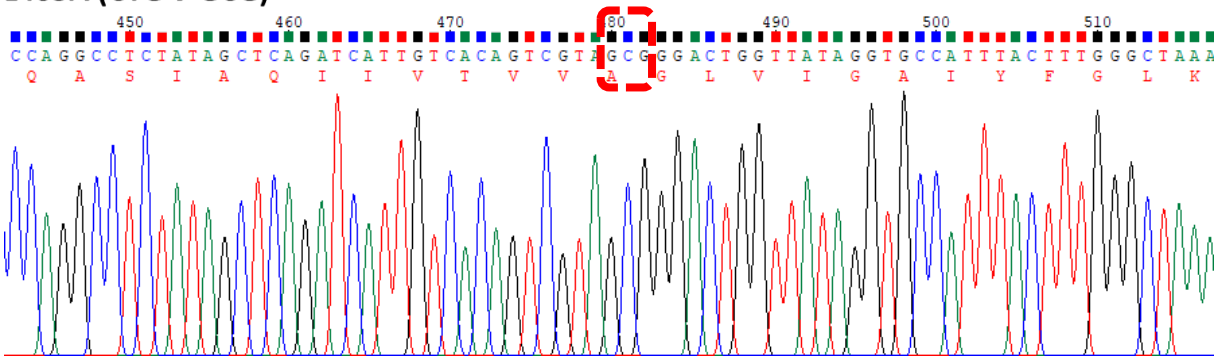
Mut1

L539A (CTT→GCT)

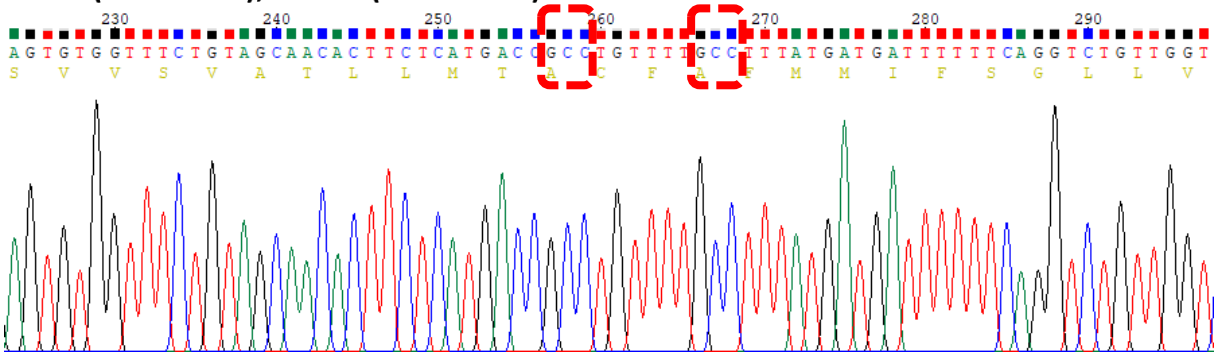


Mut2

L405A (CTG→GCG)



Mut2
I543A (ATC→GCC), V546A (GTG→GCC)



6.2 Sequence alignments

The sequence alignments of the WT construct (Query) aligned with the sequence determined by Sanger sequencing (Sbjct) is shown below for each primer.

6.2.1 T435A

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2508 bits(1358)	0.0	1373/1383(99%)	2/1383(0%)	Plus/Plus

```
Query  912   AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  75

Query  972   TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  1031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  76    TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  135

Query  1032  CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  1091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  136   CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  195

Query  1092  TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  1151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  196   TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  255

Query  1152  CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  256   CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  315

Query  1212  CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  316   CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  375

Query  1272  AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  376   AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  435

Query  1332  GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  436   GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  495

Query  1392  CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  496   CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  555

Query  1452  GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  556   GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  615

Query  1512  CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  1571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  616   CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  675

Query  1572  CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  1631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  676   CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  735

Query  1632  TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  1691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  736   TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  795

Query  1692  TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  1751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  796   TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  855

Query  1752  AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  1811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  856   AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  915

Query  1812  CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  916   CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  975

Query  1872  TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  976   TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1035

Query  1932  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1036  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1095

Query  1992  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaa  2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1096  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAA  1155

Query  2052  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct  1156  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCANACTCCAAGGT  1215

Query  2112  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAG-AAAAAGGACTAGTATAGGAA  2170
             ||||||||||||||||||||||||||||||||||||||  ||||||||||||||| ||||||
Sbjct  1216  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAANNAAAAAGGACTAGTNTAGGAA  1275

Query  2171  TGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACT  2230
             ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Sbjct  1276  TGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGNACCTACAACTGGCTTAGACT  1335

Query  2231  CAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGG-ACGAACA  2289
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |  ||||
Sbjct  1336  CAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCCGGGGANNAACA  1395

Query  2290  ATC  2292
             |||
Sbjct  1396  ATC  1398
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2362 bits(1279)	0.0	1334/1365(98%)	9/1365(0%)	Plus/Plus

```
Query  1877  GTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGTA  1936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24    GTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGTA  83

Query  1937  ATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAACT  1996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  84    ATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAACT  143

Query  1997  TACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaCGAAC  2056
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  144   TACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAACGAAC  203

Query  2057  GGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGAA  2116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  204   GGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGAA  263

Query  2117  CTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAGC  2176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  264   CTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAGC  323

Query  2177  TTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGCA  2236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  324   TTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGCA  383

Query  2237  CAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATCT  2296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  384   CAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATCT  443

Query  2297  TCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTGG  2356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  444   TCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTGG  503

Query  2357  CCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCAG  2416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  504   CCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCAG  563

Query  2417  CTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAATG  2476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  564   CTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAATG  623

Query  2477  GAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATAG  2536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  624   GAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATAG  683

Query  2537  AGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCCT  2596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  684   AGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCCT  743

Query  2597  CCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAGA  2656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  744   CCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAGA  803

Query  2657  AGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGGG  2716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  804   AGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGGG  863

Query  2717  TTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAGATCA  2776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  864   TTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAGATCA  923

Query  2777  TTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGATT  2836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  924   TTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGATT  983

Query  2837  CTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGTTTCA  2896
             |||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Sbjct  984   CTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGGCTAACCAGTGTTTCA  1043

Query  2897  GCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATACA  2956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1044  GCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATACA  1103

Query  2957  TCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTAC  3016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1104  TCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTAC  1163

Query  3017  CCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGA  3076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1164  CCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGA  1223

Query  3077  AGCCAAAGGCAGATGCCTTCTTCGTTATGAT-GTTTACCCTTATGATGGTGGCTTATTCA  3135
             |||||||||||||||||||||||||||||  ||||||||||||| ||||||||||||||
Sbjct  1224  AGCCAAAGGCAGATGCCTTCTTCGTTATGAANGTTTACCCTTATGANGGTGGCTTATTCA  1283

Query  3136  GCCAGTTCCATGG-CACTGG-CCATAGCAGC-AGGTCAGAGT-GTGGTTTCTGTAGCAAC  3191
             ||||||||||||| | |||| |||||||| | |||||| |   ||||||   ||| ||||
Sbjct  1284  GCCAGTTCCATGGNCNCTGGNCCATAGCANCCAGGTCANAANNGTGGTTNNNGTAACAAC  1343

Query  3192  ACTT-CTCATGACCATCTG-TTTTGTG-TTTATGAT-GATTTTTT  3232
                || ||   ||||||||| ||||| | |||| ||  || |||||
Sbjct  1344  NNTTNCTNNGGACCATCTGGTTTTGGGGTTTAAGAANGAATTTTT  1388
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2457 bits(1330)	0.0	1363/1384(98%)	1/1384(0%)	Plus/Plus

```
Query  2325  CTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  2384
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    CTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  74

Query  2385  TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  2444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  75    TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  134

Query  2445  CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135   CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  194

Query  2505  AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  2564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195   AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  254

Query  2565  AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  2624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255   AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  314

Query  2625  ACATCAACTTTCCGGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  2684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  315   ACATCAACTTTCCGGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  374

Query  2685  CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  2744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  375   CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  434

Query  2745  GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  2804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  435   GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  494

Query  2805  AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  2864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  495   AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  554

Query  2865  TCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGT  2924
              |||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Sbjct  555   TCTCTTCTTCCTGACGGCTAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGT  614

Query  2925  GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  2984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  615   GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  674

Query  2985  TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  3044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  675   TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  734

Query  3045  TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  3104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  735   TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  794

Query  3105  GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  3164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795   GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  854

Query  3165  AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  3224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  855   AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  914

Query  3225  GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  3284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  915   GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  974

Query  3285  GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  3344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  975   GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  1034

Query  3345  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  3404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1035  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  1094

Query  3405  TGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGAA  3464
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  1095  TGGCGAANAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAANAA  1154

Query  3465  TCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  3524
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1155  TCACGTGGCCTTGGCTTGGATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  1214

Query  3525  ATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  3584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1215  ATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  1274

Query  3585  CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCT  3644
              |||||||||||||||||||||||||||||||||||||||||||||  |||||  ||||||||
Sbjct  1275  CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCNNCCCCCNGGCCTTCCT  1334

Query  3645  TGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGC  3704
              |||||||||||| |||  ||||  |  ||||||| || |||  ||||||||||||||||
Sbjct  1335  TGACCCTGGAAGGGGCCCNTCCCCNTNNCCTTTCCAAAAAAANGGAGGAAATTGCATCGC  1394

Query  3705  ATTG  3708
              ||||
Sbjct  1395  ATTG  1398
```

Seq482

Score	Expect	Identities	Gaps	Strand
2283 bits(1236)	0.0	1285/1310(98%)	12/1310(0%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  13    GCGGNNACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  72

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  132

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  192

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  252

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  312

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  372

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  432

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  433   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  492

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  552

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  612

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  672

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  732

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGT  792

Query  3740  ggggtggggCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   GGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  852

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  853   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  912

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  972

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  1032

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTT  4039
             ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Sbjct  1033  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTT  1092

Query  4040  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  4099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  1152

Query  4100  CCATCGCCCTGATAGACGG-TTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAG  4158
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  1153  CCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAG  1212

Query  4159  TGG-ACTCTTGTTCCAAACTGG-AACAACACT-CAACCCTATCT--CGGTCTATTC-TTT  4212
             ||| ||||||| |||||||||| ||||||||| |||||     |  |||||||||| |||
Sbjct  1213  TGGAACTCTTGGTCCAAACTGGNAACAACACTTCAACCNNNNTTNNCGGTCTATTCNTTT  1272

Query  4213  TGA-TTTATAAGGGA-TTTTGGGGA-TTTCGGCC-TATTGG-TTAAAAAA  4257
             ||| ||||||||||| ||||| | |||||||| |||||| ||||||||
Sbjct  1273  TGAATTTATAAGGGAATTTTGCCNAATTTCGGCCCTATTGGGTTAAAAAA  1322
```

6.2.2 N436A/S622S

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2300 bits(1245)	0.0	1293/1323(98%)	9/1323(0%)	Plus/Plus

```
Query  912   AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  75

Query  972   TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  1031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  76    TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  135

Query  1032  CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  1091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  136   CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  195

Query  1092  TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  1151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  196   TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  255

Query  1152  CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  256   CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  315

Query  1212  CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  316   CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  375

Query  1272  AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  376   AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  435

Query  1332  GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  436   GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  495

Query  1392  CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  496   CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  555

Query  1452  GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  556   GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  615

Query  1512  CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  1571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  616   CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  675

Query  1572  CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  1631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  676   CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  735

Query  1632  TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  1691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  736   TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  795

Query  1692  TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  1751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  796   TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  855

Query  1752  AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  1811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  856   AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  915

Query  1812  CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  916   CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  975

Query  1872  TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  976   TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1035

Query  1932  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1036  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1095

Query  1992  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaa  2051
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
Sbjct  1096  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGC-ACAACTATGACGAATCATG-AAAAAA  1153

Query  2052  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  2111
             |||||||||||  |||||||||||  | ||||||||||| ||||||||||||||||||||
Sbjct  1154  CGAACGGATTANNAGGGTCATTCAGAANTTAGGTCTGGAT-AAGTGGCAGACTCCAAGGT  1212

Query  2112  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAAT  2171
              |||||||  |||||||| |   | ||||||||||||||||||||||||||  ||||||
Sbjct  1213  -GGAACTCN-TTTATCCGNGNN-TNTCTGGAGGAGAAAGAAAAAGGACTAGTNNAGGAAT  1269

Query  2172  GGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTC  2231
             |||||||||  ||||||||||    ||||||| |||||| | | || |||||||| | ||
Sbjct  1270  GGAGCTTATNNCTGATCCTTCCNNNTTGTTCT-GGATGACCNTNCA-CTGGCTTACA-TC  1326

Query  2232  AAG  2234
             |||
Sbjct  1327  AAG  1329
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2362 bits(1279)	0.0	1318/1340(98%)	6/1340(0%)	Plus/Plus

```
Query  1873  CCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAA  1932
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20    CCNAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAA  79

Query  1933  TGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAA  1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80    TGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAA  139

Query  1993  AACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaC  2052
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140   AACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAAC  199

Query  2053  GAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTT  2112
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200   GAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTT  259

Query  2113  GGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATG  2172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260   GGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATG  319

Query  2173  GAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCA  2232
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320   GAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCA  379

Query  2233  AGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATC  2292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  380   AGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATC  439

Query  2293  ATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTA  2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  440   ATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTA  499

Query  2353  TTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAA  2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  500   TTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAA  559

Query  2413  TCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATT  2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  560   TCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATT  619

Query  2473  AATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATC  2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  620   AATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATC  679

Query  2533  ATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAAC  2592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  680   ATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAAC  739

Query  2593  TCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAG  2652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  740   TCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAG  799

Query  2653  AAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGA  2712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  800   AAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGA  859

Query  2713  TGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAG  2772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  860   TGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAG  919

Query  2773  ATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAAT  2832
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  920   ATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAAT  979

Query  2833  GATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGT  2892
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Sbjct  980   GATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCGCCCAGTGT  1039

Query  2893  TTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAA  2952
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1040  TTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAA  1099

Query  2953  TACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTA  3012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1100  TACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTA  1159

Query  3013  TTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGA  3072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1160  TTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGA  1219

Query  3073  TTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGT-GGCTTA  3131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||  ||||||
Sbjct  1220  TTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGAAGGGNGGCTTA  1279

Query  3132  TTCAGCCAGTTCCAT-GGCACTGGCCATAGCAGC-AGGTCAGAGTG-TGGTTTCTG-TAG  3187
             ||||||||||||||| ||||||||||||  | | ||||||||  | ||||| | | ||
Sbjct  1280  TTCAGCCAGTTCCATNGGCACTGGCCATANNAACCAGGTCAGANNGGTGGTTNCNGGTAA  1339

Query  3188  CAACACTT-CTCATGACCAT  3206
             |||  ||| |||| ||||||
Sbjct  1340  CAANCCTTTCTCAGGACCAT  1359
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2492 bits(1349)	0.0	1396/1427(98%)	5/1427(0%)	Plus/Plus

```
Query  2324  TCTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  2383
             |||||                                                        |
Sbjct  14    TCTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  72

Query  2384  CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  132

Query  2444  ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  192

Query  2504  GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  252

Query  2564  TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  312

Query  2624  TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  2683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  372

Query  2684  ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  2743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  432

Query  2744  TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  2803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  433   TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  492

Query  2804  TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  2863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  552

Query  2864  TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  2923
             ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Sbjct  553   TTCTCTTCTTCCTGACGACCGCCCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  612

Query  2924  TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  2983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  672

Query  2984  ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  3043
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  732

Query  3044  TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  3103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  792

Query  3104  TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  3163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  852

Query  3164  CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGA  3223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  853   CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGA  912

Query  3224  TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  3283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  972

Query  3284  AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  3343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  1032

Query  3344  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  3403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1033  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  1092

Query  3404  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGA  3463
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||| |
Sbjct  1093  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCGCCCTGGGGCTTGTGGAANA  1152

Query  3464  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  3523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1153  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  1212

Query  3524  TATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  3583
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1213  TATTTCTTAAAAANTATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  1272

Query  3584  GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC  3643
             ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1273  GCCTCGACTGTGCCTTNNAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC  1332

Query  3644  TTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAAT-GAGGAAATTGCATC  3702
             ||||||||||||||||||    ||    |  |||||||  |||||||  ||||||||||||
Sbjct  1333  TTGACCCTGGAAGGTGCCNNNCCNN-TNNCCTTTCC-AATAAAANGGAGGAAATTGCATC  1390

Query  3703  GCATTGTCTG-AGTAGGTGTCATTCTATTCTgggggtggggtgggg  3748
             |||||| | | | |||||  ||||   ||| ||||||  |||| ||||
Sbjct  1391  GCATTGGCCGNANTAGGTNNCATTNNNTTCGGGGGGGGGGGGGGGG  1437
```

Seq482

Score	Expect	Identities	Gaps	Strand
2052 bits(1111)	0.0	1143/1166(98%)	2/1166(0%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11    GCGGNNACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  70

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  71    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  130

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  131   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  190

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  191   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  250

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  251   TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  310

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  311   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  370

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  371   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  430

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  431   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  490

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  491   TCTCGCCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  550

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  551   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  610

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  611   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  670

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  671   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  730

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  731   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGT  790

Query  3740  ggggtgggCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  791   GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  850

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             |||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||
Sbjct  851   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGNGCTCTAGNGGGTATCCC  910

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  911   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  970

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             |||||||||||||||||||| |||||||||||||||||||| |||||||||||||||||
Sbjct  971   GCTACACTTGCCAGCGCCCTANCGCCCGCTCCTTTCGCTTTCTTCNCTTCCTTTCTCGCC  1030

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGG-CATCCCTTTAGGGTTCCGATT  4038
             |||||||||||||||||||| ||||||||||||||||   |||||||||| ||||||||||
Sbjct  1031  ACGTTCGCCGGCTTTCCCCGTCGAGCTCTAAATCGGGGGNCTCCCTTTAGGNTTCCGATT  1090

Query  4039  TAGTGCTTTACGGCACCTCGACCCCAAAAA-ACTTGATTAGGGTGATGGTTCACGTAGTG  4097
             |||||||||||||||||||| ||||| |||| ||||||||||||||| |||||||||| ||
Sbjct  1091  TAGTGCTTTACGGCACCTCNACCCCNAAAANACTTGATTAGGGTGAAGGTTCACGTANTG  1150

Query  4098  GGCCATCGCCCTGATAGACGGTTTTT  4123
             |||||||  |||||   |||| ||||
Sbjct  1151  GGCCATCNNCCTGAANNACGGCTTTT  1176
```

6.2.3 F439A

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2558 bits(1385)	0.0	1414/1436(98%)	3/1436(0%)	Plus/Plus

```
Query  907   ACTTAAGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGA  966
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15    ACTTNNGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGA  74

Query  967   AGCGGTGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACC  1026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  75    AGCGGTGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACC  134

Query  1027  ATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAA  1086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135   ATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAA  194

Query  1087  CTGTCTGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCC  1146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195   CTGTCTGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCC  254

Query  1147  GCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATG  1206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255   GCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATG  314

Query  1207  CAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCT  1266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  315   CAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCT  374

Query  1267  GTGCCAGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGG  1326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  375   GTGCCAGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGG  434

Query  1327  AAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTG  1386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  435   AAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTG  494

Query  1387  GCCGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCC  1446
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  495   GCCGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCC  554

Query  1447  ATTCTGATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGC  1506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  555   ATTCTGATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGC  614

Query  1507  GGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGG  1566
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  615   GGGCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGG  674

Query  1567  CTGGGCGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAAC  1626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  675   CTGGGCGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAAC  734

Query  1627  ACCAATGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTG  1686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  735   ACCAATGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTG  794

Query  1687  TTAAGTTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGA  1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795   TTAAGTTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGA  854

Query  1747  AAACCAGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAAC  1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  855   AAACCAGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAAC  914

Query  1807  GCCATCCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGG  1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  915   GCCATCCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGG  974

Query  1867  AAAGATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAAC  1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  975   AAAGATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAAC  1034

Query  1927  TTCAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTG  1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1035  TTCAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTG  1094

Query  1987  AGAGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaa  2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1095  AGAGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAA  1154

Query  2047  aaaaaCGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCC  2106
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1155  AAAAACGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCC  1214

Query  2107  AAGGTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAA-GGACTAGTAT  2165
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||||||
Sbjct  1215  AAGGTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAANGAAAAAGGGACTAGTAT  1274

Query  2166  AGGAATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTT  2225
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  1275  AGGAATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAACCTACAACTGGCTT  1334

Query  2226  AGACTCAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACG  2285
             ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1335  AGACTCAAGCNNAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGANN  1394

Query  2286  AACAATCATCTTCT-CCATTCATCAGCCTCG-ATATTCCATCTTCAAGTTGTTTGA  2339
             |||||||| |||   |||||||||  || |  |  ||||||||||||||  ||||
Sbjct  1395  AACAATCANCTTNNNCCATTCATCNNCCNCCNANNTTCCATCTTCAAGTTNNTTGA  1450
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2386 bits(1292)	0.0	1333/1353(99%)	8/1353(0%)	Plus/Plus

```
Query  1870  GATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTC  1929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17    GATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTC  76

Query  1930  AAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGA  1989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  77    AAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGA  136

Query  1990  GAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaa  2049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  137   GAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAA  196

Query  2050  aaCGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAG  2109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  197   AACGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAG  256

Query  2110  GTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGA  2169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  257   GTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGA  316

Query  2170  ATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGAC  2229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  317   ATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGAC  376

Query  2230  TCAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACA  2289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  377   TCAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACA  436

Query  2290  ATCATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACC  2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  437   ATCATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACC  496

Query  2350  TTATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTT  2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  497   TTATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTT  556

Query  2410  GAATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATC  2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  557   GAATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATC  616

Query  2470  ATTAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAG  2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  617   ATTAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAG  676

Query  2530  ATCATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTC  2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  677   ATCATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTC  736

Query  2590  AACTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAG  2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  737   AACTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAG  796

Query  2650  AAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTC  2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  797   AAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTC  856

Query  2710  AGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCT  2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  857   AGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCT  916

Query  2770  CAGATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAA  2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  917   CAGATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAA  976

Query  2830  AATGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAG  2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  977   AATGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAG  1036

Query  2890  TGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACAT  2949
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1037  TGTGCCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACAT  1096

Query  2950  GAATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGAT  3009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1097  GAATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGAT  1156

Query  3010  TTATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTA  3069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1157  TTATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTA  1216

Query  3070  GGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGAT-GGTGGC  3128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Sbjct  1217  GGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGAAGGGTGGC  1276

Query  3129  TTATTCAGCCAGTTCCATGG-CACTGG-CCATAGC-AGCAGGTCAG-AGT-GTGGTTTCT  3183
             |||||||||||||||||||| |||||| ||||||| | |||||| | |   |||||||||
Sbjct  1277  TTATTCAGCCAGTTCCATGGGCACTGGNCCATAGCNACCAGGTCNGNAANNGTGGTTTCT  1336

Query  3184  GTAGCAACACTTCTC-ATGACCATCTG-TTTTG  3214
             ||| ||||||||  | | ||||||| | |||||
Sbjct  1337  GTAACAACACTTNCCNAGGACCATCCGGTTTTG  1369
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2396 bits(1297)	0.0	1323/1338(99%)	2/1338(0%)	Plus/Plus

```
Query  2325  CTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  2384
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    CTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  74

Query  2385  TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  2444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  75    TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  134

Query  2445  CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135   CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  194

Query  2505  AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  2564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195   AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  254

Query  2565  AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  2624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255   AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  314

Query  2625  ACATCAACTTTCCGGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  2684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  315   ACATCAACTTTCCGGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  374

Query  2685  CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  2744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  375   CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  434

Query  2745  GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  2804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  435   GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  494

Query  2805  AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  2864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  495   AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  554

Query  2865  TCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGT  2924
              ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Sbjct  555   TCTCTTCTTCCTGACGACCAACCAGTGTGCCAGCAGTGTTTCAGCCGTGGAACTCTTTGT  614

Query  2925  GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  2984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  615   GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  674

Query  2985  TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  3044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  675   TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  734

Query  3045  TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  3104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  735   TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  794

Query  3105  GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  3164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795   GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  854

Query  3165  AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  3224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  855   AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  914

Query  3225  GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  3284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  915   GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  974

Query  3285  GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  3344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  975   GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  1034

Query  3345  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  3404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1035  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  1094

Query  3405  TGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGAA  3464
              |||||||  ||||||||||||||||||||||||||||||||||||||||||||||  ||
Sbjct  1095  TGGCGAANAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAANAA  1154

Query  3465  TCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  3524
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1155  TCACGTGGCCTTGGCTTGGATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  1214

Query  3525  ATTTCTTAAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  3584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1215  ATTTCTTAAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  1274

Query  3585  CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCC-TCCCCCGTGCCTTCC  3643
              |  | |||  |||||||  |||||||||||||||||||||||||||  |||||||  |||||||
Sbjct  1275  CTCCAACTNTGCCTTCNAGTTGCCAGCCATCTGTTGTTTGCCCCNTCCCCCGGGCCTTCC  1334

Query  3644  TTGACCCTGGAAGGTGCC  3661
              ||||||  |||||||  |||
Sbjct  1335  TTGACCNTGGAAGGGGCC  1352
```

Seq482

Score	Expect	Identities	Gaps	Strand
2300 bits(1245)	0.0	1301/1330(98%)	16/1330(1%)	Plus/Plus

```
Query  2966  ACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGA  3025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20    ACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGA  79

Query  3026  TGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGG  3085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80    TGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGG  139

Query  3086  CAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCA  3145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140   CAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCA  199

Query  3146  TGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCA  3205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200   TGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCA  259

Query  3206  TCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCAT  3265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260   TCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCAT  319

Query  3266  CTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGC  3325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320   CTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGC  379

Query  3326  ATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTT  3385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  380   ATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTT  439

Query  3386  GTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCAC  3445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  440   GTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCAC  499

Query  3446  CCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAA  3505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  500   CCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAA  559

Query  3506  TTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCC  3565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  560   TTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCC  619

Query  3566  GTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC  3625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  620   GTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC  679

Query  3626  CCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA  3685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  680   CCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA  739

Query  3686  AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggtggggtg  3745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  740   AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTG  799

Query  3746  gggCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTG  3805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  800   GGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTG  859

Query  3806  GGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCG  3865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  860   GGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCG  919

Query  3866  CCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA  3925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  920   CCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA  979

Query  3926  CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC  3985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  980   CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC  1039

Query  3986  GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTTAGTGCT  4045
              ||||||||||||||||||||||||||||||||||  || ||||||||||||||||||||||
Sbjct  1040  GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCNCTTTAGGGTTCCGATTTAGTGCT  1099

Query  4046  TTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCG  4105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1100  TTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCG  1159

Query  4106  CCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTC  4165
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| ||||||
Sbjct  1160  CCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCANGTTCTTTAATAGNGGACTC  1219

Query  4166  TTGTTCCAAACTGG-AACAACAC-TCAACCCTAT--CTCGG-TCTATTC-TTTTGA-TTT  4218
              ||||||||||||| || ||||| |||||||||    ||||| ||||||| |||||| |||
Sbjct  1220  TTGTTCCAAACTGGNAANAACACCTCAACCCTANNNCTCGGGTCTATTCCTTTTGAATTT  1279

Query  4219  ATAAGGGA-TTTTGGGGA-TTTC-GGCCTATTGG-TTAAAAAAT-G-AG-CTGATTT-AA  4270
              | |||||| |||||   | |||| |||||||||| ||||||||  | |  ||||||| ||
Sbjct  1280  ANAAGGGAATTTTGNCNAATTTCCGGCCTATTGGGTTAAAAAANNGNANNCTGATTTTAA  1339

Query  4271  CAAAAA-TTT  4279
              |||||| |||
Sbjct  1340  CAAAAANTTT  1349
```

6.2.4 S440W

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2497 bits(1352)	0.0	1368/1380(99%)	1/1380(0%)	Plus/Plus

```
Query  914   CTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGGTG  973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20    CTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGGTG  79

Query  974   GAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGACA  1033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80    GAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGACA  139

Query  1034  AAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTCTG  1093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140   AAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTCTG  199

Query  1094  GGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGACG  1153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200   GGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGACG  259

Query  1154  CCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCA  1213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260   CCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCA  319

Query  1214  CCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAG  1273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320   CCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAG  379

Query  1274  CCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGC  1333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  380   CCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGC  439

Query  1334  TGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGGCA  1393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  440   TGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGGCA  499

Query  1394  ATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCTGA  1453
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  500   ATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCTGA  559

Query  1454  TCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCTCG  1513
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  560   TCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCTCG  619

Query  1514  CCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGCG  1573
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  620   CCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGCG  679

Query  1574  AATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAATG  1633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  680   AATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAATG  739

Query  1634  GCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAGTT  1693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  740   GCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAGTT  799

Query  1694  TTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACCAG  1753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  800   TTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACCAG  859

Query  1754  TTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCATCC  1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  860   TTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCATCC  919

Query  1814  TGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGATC  1873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  920   TGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGATC  979

Query  1874  CAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAAT  1933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  980   CAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAAT  1039

Query  1934  GTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAA  1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1040  GTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAA  1099

Query  1994  ACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaCG  2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1100  ACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAACG  1159

Query  2054  AACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTG  2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1160  AACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTG  1219

Query  2114  GAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGG  2173
             |||||||||||||||||| ||||||||||||  ||||||||||||||||||||||||||
Sbjct  1220  GAACTCAGTTTATCCGTGGTNTGTCTGGAGGAGAANNAAAAAGGACTAGTATAGGAATGG  1279

Query  2174  AGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAG-ACTCA  2232
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||   |||||
Sbjct  1280  AGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAACCTACAACTGGCTTATAACTCA  1339

Query  2233  AGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATC  2292
             |||  ||||||||||||||||||||||||||||||||||||| ||||||  || ||||
Sbjct  1340  AGCNNAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAANCAGGGANNAANAATC  1399
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2392 bits(1295)	0.0	1358/1391(98%)	11/1391(0%)	Plus/Plus

```
Query  1869  AGATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTT  1928
             ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    AGATCC-AGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTT  74

Query  1929  CAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAG  1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  75    CAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAG  134

Query  1989  AGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaa  2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135   AGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAA  194

Query  2049  aaaCGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAA  2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195   AAACGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAA  254

Query  2109  GGTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGG  2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255   GGTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGG  314

Query  2169  AATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGA  2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  315   AATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGA  374

Query  2229  CTCAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAAC  2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  375   CTCAAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAAC  434

Query  2289  AATCATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCAC  2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  435   AATCATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCAC  494

Query  2349  CTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTT  2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  495   CTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTT  554

Query  2409  TGAATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACAT  2468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  555   TGAATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACAT  614

Query  2469  CATTAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGA  2528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  615   CATTAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGA  674

Query  2529  GATCATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGT  2588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  675   GATCATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGT  734

Query  2589  CAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAA  2648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  735   CAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAA  794

Query  2649  GAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACT  2708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  795   GAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACT  854

Query  2709  CAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGC  2768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  855   CAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGC  914

Query  2769  TCAGATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAA  2828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  915   TCAGATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAA  974

Query  2829  AAATGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCA  2888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  975   AAATGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCA  1034

Query  2889  GTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACA  2948
              ||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1035  GTGTTTTTGGAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACA  1094

Query  2949  TGAATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGA  3008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1095  TGAATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGA  1154

Query  3009  TTTATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTT  3068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1155  TTTATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTT  1214

Query  3069  AGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGT-GG  3127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||  ||
Sbjct  1215  AGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGAAGGGNGG  1274

Query  3128  CTTATTCAGCCAGTTCCATGGCACTGG-CCATAGCAGCAGGTCAGAGT-GTGGTTTCTGT  3185
              ||||||||||||||||||||||||  |||| ||||||||  |||||||||   ||||||||||
Sbjct  1275  CTTATTCAGCCAGTTCCATGGCNCTGGNCCATAGCANCAGGTCAGANNNGTGGTTTCTGG  1334

Query  3186  AGCAACACTTCTC-ATGACCATCTG-TTTTGTGTTT-ATG-ATGA-TTTTTTCAGG-TCT  3239
              | |||  |||    ||||||||||| ||||| | || |   | || ||||||||| ||
Sbjct  1335  ANCAANCCTTTNNNATGACCATCTGGTTTTGGGGTTNAANNAAGAATTTTTTCAGGGTCG  1394

Query  3240  GTTGG-TCAAT  3249
              ||||| |||||
Sbjct  1395  GTTGGGTCAAT  1405
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2466 bits(1335)	0.0	1372/1397(98%)	4/1397(0%)	Plus/Plus

```
Query  2325  CTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  2384
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15    CTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC  73

Query  2385  TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  2444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  74    TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA  133

Query  2445  CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  134   CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG  193

Query  2505  AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  2564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  194   AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT  253

Query  2565  AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  2624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  254   AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT  313

Query  2625  ACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  2684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  314   ACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA  373

Query  2685  CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  2744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  374   CACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGCT  433

Query  2745  GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  2804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  434   GGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT  493

Query  2805  AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  2864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  494   AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT  553

Query  2865  TCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGT  2924
              |||||||||||||||||||||||||||||  | ||||||||||||||||||||||||||||
Sbjct  554   TCTCTTCTTCCTGACGACCAACCAGTGTTTTTGGAGTGTTTCAGCCGTGGAACTCTTTGT  613

Query  2925  GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  2984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  614   GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTTA  673

Query  2985  TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  3044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  674   TTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT  733

Query  3045  TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  3104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  734   TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT  793

Query  3105  GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  3164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  794   GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAGC  853

Query  3165  AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  3224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  854   AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGAT  913

Query  3225  GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  3284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  914   GATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTCA  973

Query  3285  GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  3344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  974   GTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGACA  1033

Query  3345  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  3404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1034  AAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTAC  1093

Query  3405  TGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGAA  3464
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1094  TGGCGAANAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGAA  1153

Query  3465  TCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  3524
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1154  TCACGTGGCCTTGGCTTGGATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGTT  1213

Query  3525  ATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  3584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1214  ATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCAG  1273

Query  3585  CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT-GTTTGCCCCTCCCCCGTGCCTTCC  3643
              |||||||||||||||| |||||||||||||  | | |||||||||||||||| |||||||
Sbjct  1274  CCTCGACTGTGCCTTCNAGTTGCCAGCCATNNGNTNGTTTGCCCCTCCCCCGGGCCTTCC  1333

Query  3644  TTGACCCTGGAAGGTGCCACTCCCACTGT-CCTTTCCTAATAAAATGAGGAAATTGCATC  3702
              |||||||||||||| |||  | ||| ||  |||||||| ||||||    ||||| |||||||
Sbjct  1334  TTGACCCTGGAAGGGGCCNNTNCCANTGNNCCTTTCC-AATAAANNNAGGAANTTGCATC  1392

Query  3703  GCATTGTCTGAGTAGGT  3719
              ||||||||||| |||||
Sbjct  1393  GCATTGTCTGANTAGGT  1409
```

Seq482

Score	Expect	Identities	Gaps	Strand
2287 bits(1238)	0.0	1303/1337(97%)	17/1337(1%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  12    GCGGNNACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  71

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  72    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  131

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  191

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  192   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  251

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  252   TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  311

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  312   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  371

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  372   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  431

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  432   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  491

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  492   TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  551

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  552   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  611

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  612   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  671

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  731

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  732   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGT  791

Query  3740  ggggtgggg CAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  792   GGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  851

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  852   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  911

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  912   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  971

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  972   GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  1031

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTT  4039
             ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Sbjct  1032  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTT  1091

Query  4040  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  4099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1092  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  1151

Query  4100  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGG-AGTCCACGTTCTTTAATAG  4158
             ||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  1152  CCATCGCCCTGATANACGGTTTTTCGCCCTTTGACGTTGGAAGTCCACGTTCTTTAATAG  1211

Query  4159  TGG-ACTCTTGTTCCAAACTGGAA-CAACACT-CAACCCTATCT-CGG-TCTATTCTTTT  4213
             ||| ||||||||||||| |||||| |||    | ||||||||||  ||| ||||||||||
Sbjct  1212  TGGAACTCTTGTTCCAANCTGGAAACAANNNTTCAACCCTATCCNCGGNTCTATTCTTTT  1271

Query  4214  -GATTTATAA-GGGA-TTTTGGGGA-TTTC-GGCCTATTGG-TTAAAAAA-TGAGC-TG-  4264
              |||||  || |||| |||||    | |||| |||||||||| ||||||| ||   | ||
Sbjct  1272  TGATTTNNAAAGGGAATTTTGCCNAATTTCCGGCCTATTGGGTTAAAAAAATGNNCCTGG  1331

Query  4265  ATTT-AACAAAAA-TTT  4279
             |||| |||||||| |||
Sbjct  1332  ATTTTAACAAAAANTTT  1348
```

6.2.5 M549E

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2499 bits(1353)	0.0	1382/1400(99%)	4/1400(0%)	Plus/Plus

```
Query  909   TTAAGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAG  968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14    TTAAGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAG  73

Query  969   CGGTGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCAT  1028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  74    CGGTGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCAT  133

Query  1029  GGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACT  1088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  134   GGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACT  193

Query  1089  GTCTGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGC  1148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  194   GTCTGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGC  253

Query  1149  CGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCA  1208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  254   CGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCA  313

Query  1209  GGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGT  1268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  314   GGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGT  373

Query  1269  GCCAGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAA  1328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  374   GCCAGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAA  433

Query  1329  ACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGC  1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  434   ACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGC  493

Query  1389  CGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCAT  1448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  494   CGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCAT  553

Query  1449  TCTGATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGG  1508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  554   TCTGATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGG  613

Query  1509  GCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCT  1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  614   GCTCGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCT  673

Query  1569  GGGCGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACAC  1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  674   GGGCGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACAC  733

Query  1629  CAATGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTT  1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  734   CAATGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTT  793

Query  1689  AAGTTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAA  1748
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  794   AAGTTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAA  853

Query  1749  ACCAGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGC  1808
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  854   ACCAGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGC  913

Query  1809  CATCCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAA  1868
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  914   CATCCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAA  973

Query  1869  AGATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTT  1928
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  974   AGATCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTT  1033

Query  1929  CAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAG  1988
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1034  CAAATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAG  1093

Query  1989  AGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaa  2048
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1094  AGAAAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAA  1153

Query  2049  aaaCGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAA  2108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1154  AAACGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAA  1213

Query  2109  GGTTGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGG  2168
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  1214  GGTTGGAACTCAGTTTATCCGTGGTGNGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGG  1273

Query  2169  AATGGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAG-  2227
             ||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Sbjct  1274  AATGGAGCTTATCACTGATCCTTCCATCTTGGTCTTGGATGAACCTACAACTGGCTTAAA  1333

Query  2228  ACTCAAGCACAGCAAAT-GCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGA  2286
             |||||||||||||||  |||||||||| |  |||||||||||||||||| | |||||| |
Sbjct  1334  ACTCAAGCACAGCAAANGGCTGTCCTTTNGNNCCTGAAAAGGATGTCTAACCNGGGACAA  1393

Query  2287  AC-AATCATC-TTCTCCATT  2304
             |  ||||| | || ||||||
Sbjct  1394  ANNAATCANCNTTNTCCATT  1413
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2348 bits(1271)	0.0	1303/1320(99%)	7/1320(0%)	Plus/Plus

```
Query  1872  TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1931
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19    TCCNAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  78

Query  1932  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  79    ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  138

Query  1992  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaa  2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  139   AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAA  198

Query  2052  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  199   CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  258

Query  2112  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAAT  2171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  259   TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAAT  318

Query  2172  GGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTC  2231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  319   GGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTC  378

Query  2232  AAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAAT  2291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  379   AAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAAT  438

Query  2292  CATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTT  2351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  439   CATCTTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTT  498

Query  2352  ATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGA  2411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  499   ATTGGCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGA  558

Query  2412  ATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCAT  2471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  559   ATCAGCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCAT  618

Query  2472  TAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGAT  2531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  619   TAATGGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGAT  678

Query  2532  CATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAA  2591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  679   CATAGAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAA  738

Query  2592  CTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGGTGAGAAGAA  2651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  739   CTCCTCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGGTGAGAAGAA  798

Query  2652  GAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAG  2711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  799   GAAGAAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAG  858

Query  2712  ATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCA  2771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  859   ATGGGTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCA  918

Query  2772  GATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAA  2831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  919   GATCATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAA  978

Query  2832  TGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTG  2891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  979   TGATTCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTG  1038

Query  2892  TTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGA  2951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1039  TTTCAGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGA  1098

Query  2952  ATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTT  3011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1099  ATACATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTT  1158

Query  3012  ATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGG  3071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1159  ATTACCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGG  1218

Query  3072  ATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGAT-GTTTACCCTTATGAT-GGTGGCT  3129
             |||||||||||||||||||||||||||||||||||   |||||||||||||  |||||||
Sbjct  1219  ATTGAAGCCAAAGGCAGATGCCTTCTTCGTTATGAAGGTTTACCCTTATGAAGGGTGGCT  1278

Query  3130  TATTCAGCCAGTTCCATGGC-ACTGG-CCATAGCAGC-AGGTCAGAGT--GTGGTTTCTG  3184
             |||||  |||||||||||||   |||||  |||||  |   |  |||||||     ||||||||||
Sbjct  1279  TATTCNGCCAGTTCCATGGGNACTGGGCCATANCNNCCAGGTCAGAANNGGTGGTTTCTG  1338
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2348 bits(1271)	0.0	1327/1363(97%)	7/1363(0%)	Plus/Plus

```
Query  2324  TCTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  2383
             |||||  || ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14    TCTTC-AGNTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  72

Query  2384  CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  2443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  132

Query  2444  ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  2503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  192

Query  2504  GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  2563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  252

Query  2564  TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  2623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  312

Query  2624  TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  2683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  372

Query  2684  ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  2743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  432

Query  2744  TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  2803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  433   TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  492

Query  2804  TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  2863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  552

Query  2864  TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  2923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  612

Query  2924  TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  2983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  672

Query  2984  ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  3043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  732

Query  3044  TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  3103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  792

Query  3104  TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  3163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  852

Query  3164  CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGA  3223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  853   CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTCATGG  912

Query  3224  TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  3283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   AGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  972

Query  3284  AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  3343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  1032

Query  3344  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  3403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1033  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  1092

Query  3404  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGA  3463
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  1093  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAANA  1152

Query  3464  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  3523
              |||||||||||||||||| ||||||||||||||||  ||||||||||||||||||||||
Sbjct  1153  ATCACGTGGCCTTGGCTTGGATGATTGTTATTTTCCTC-CAATTGCCTACCTGAAATTGT  1211

Query  3524  TATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  3583
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1212  TATTTCTTAAAAAANNTTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  1271

Query  3584  GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC  3643
              ||||| |    |||||  | ||||| |||| |||||||| |||   |||| ||| |||||
Sbjct  1272  GCCTCAANNNGGCCTTNNA-TTGCCNGCCANCTGTTGTTNGCCNNNCCCC-GTG-CTTCC  1328

Query  3644  TTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA  3686
              ||||||  |||||| ||| |  ||   ||||||||| ||||||
Sbjct  1329  TTGACCNNGGAAGGGGCC-CNNCCCNGGTCCTTTCC-AATAAA  1369
```

Seq482

Score	Expect	Identities	Gaps	Strand
2375 bits(1286)	0.0	1309/1323(99%)	0/1323(0%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  13    GCGGANACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  72

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  132

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  192

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  252

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             |||||||||||||||||||| |||  ||||||||||||||||||||||||||||||||||
Sbjct  253   TGACCATCTGTTTTGTGTTCATGGAGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  312

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  372

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  432

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  433   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  492

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  552

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  612

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  672

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  732

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGT  792

Query  3740  ggggtgggggCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  852

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  853   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  912

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  972

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  1032

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTT  4039
             ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Sbjct  1033  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTT  1092

Query  4040  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  4099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  1093  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAATGGG  1152

Query  4100  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  4159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1153  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  1212

Query  4160  GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA  4219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1213  GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTN  1272

Query  4220  TAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTT  4279
              ||||||||||    |||||||||||||||||||||||| ||||||||||||||||||
Sbjct  1273  NAAGGGATTTTGCCNATTTCGGCCTATTGGTTAAAAAATGACCTGATTTAACAAAAATTN  1332

Query  4280  AAC  4282
             |||
Sbjct  1333  AAC  1335
```

6.2.6 A397S/V401A/L539A – Mut1

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2519 bits(1364)	0.0	1377/1387(99%)	1/1387(0%)	Plus/Plus

```
Query  912   AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17    AGCTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGG  76

Query  972   TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  1031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  77    TGGAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGA  136

Query  1032  CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  1091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  137   CAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTC  196

Query  1092  TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  1151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  197   TGGGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGA  256

Query  1152  CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  257   CGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGC  316

Query  1212  CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  317   CACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCC  376

Query  1272  AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  377   AGCCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACT  436

Query  1332  GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  437   GCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGG  496

Query  1392  CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  497   CAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCT  556

Query  1452  GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  557   GATCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCT  616

Query  1512  CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  1571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  617   CGCCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG  676

Query  1572  CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  1631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  677   CGAATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAA  736

Query  1632  TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  1691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  737   TGGCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAG  796

Query  1692  TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  1751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  797   TTTTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACC  856

Query  1752  AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  1811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  857   AGTTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCAT  916

Query  1812  CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  917   CCTGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGA  976

Query  1872  TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  977   TCCAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAA  1036

Query  1932  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1037  ATGTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGA  1096

Query  1992  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaa  2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1097  AAACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAA  1156

Query  2052  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1157  CGAACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGT  1216

Query  2112  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAAT  2171
             ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Sbjct  1217  TGGAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAANNAAAAAGGACTAGTATAGGAAT  1276

Query  2172  GGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTC  2231
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1277  GGAGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAACCTACAACTGGCTTAGACTC  1336

Query  2232  AAGCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAG-CAGGGACGAACAA  2290
             |||  ||||||||||||||||||||||||||||||||||||||||  ||||||| || ||
Sbjct  1337  AAGNNCAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAANCCAGGGACNAANAA  1396

Query  2291  TCATCTT  2297
             ||| |||
Sbjct  1397  TCANCTT  1403
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2333 bits(1263)	0.0	1289/1302(99%)	5/1302(0%)	Plus/Plus

```
Query  1876  AGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGT  1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24    AGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGT  83

Query  1936  AATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAAC  1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  84    AATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAAC  143

Query  1996  TTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaCGAA  2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  144   TTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAACGAA  203

Query  2056  CGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGA  2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  204   CGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGA  263

Query  2116  ACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAG  2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  264   ACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAG  323

Query  2176  CTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGC  2235
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  324   CTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGC  383

Query  2236  ACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATC  2295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  384   ACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATC  443

Query  2296  TTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTG  2355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  444   TTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTG  503

Query  2356  GCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCA  2415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  504   GCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCA  563

Query  2416  GCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAAT  2475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  564   GCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAAT  623

Query  2476  GGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATA  2535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  624   GGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATA  683

Query  2536  GAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAAATTAGCGGAGATTTATGTCAACTCC  2595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  684   GAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAAATTAGCGGAGATTTATGTCAACTCC  743

Query  2596  TCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAG  2655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  744   TCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAG  803

Query  2656  AAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGG  2715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  804   AAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGG  863

Query  2716  GTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAGATC  2775
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  864   GTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATATCTCAGATC  923

Query  2776  ATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGAT  2835
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  924   ATTGCCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGAT  983

Query  2836  TCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGTTTC  2895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  984   TCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGTTTC  1043

Query  2896  AGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATAC  2955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1044  AGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATAC  1103

Query  2956  ATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTA  3015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1104  ATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTA  1163

Query  3016  CCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTG  3075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1164  CCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTG  1223

Query  3076  AAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCC-TTATGAT-GGTGGCTTATT  3133
             |||||||||||||||||||||||||||||| |||||||| ||||||  |||||||| ||
Sbjct  1224  AAGCCAAAGGCAGATGCCTTCTTCGTTATGAGGTTTACCCCTTATGAAGGGTGGCTTNTT  1283

Query  3134  CAGCCAGTTCCATGG-CACT-GGCCATAGCAGC-AGGTCAGA  3172
             ||||||||||||||| | || |||||||||| | ||||| ||
Sbjct  1284  CAGCCAGTTCCATGGNCNCTNGGCCATAGCAACCAGGTCNGA  1325
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2410 bits(1305)	0.0	1362/1394(98%)	10/1394(0%)	Plus/Plus

```
Query  2324  TCTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  2383
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14    TCTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  72

Query  2384  CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  132

Query  2444  ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  192

Query  2504  GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  252

Query  2564  TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  312

Query  2624  TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  2683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  372

Query  2684  ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  2743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  432

Query  2744  TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  2803
             |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Sbjct  433   TGGGTAATCCCCAGGCCTCTATATCTCAGATCATTGCCACAGTCGTACTGGGACTGGTTA  492

Query  2804  TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  2863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  552

Query  2864  TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  2923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  612

Query  2924  TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  2983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  672

Query  2984  ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  3043
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  732

Query  3044  TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  3103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  792

Query  3104  TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  3163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  852

Query  3164  CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGA  3223
             |||||||||||||||||||||||||||||  ||| ||||||||||||||||||||||||
Sbjct  853   CAGGTCAGAGTGTGGTTTCTGTAGCAACAGCTCTTATGACCATCTGTTTTGTGTTTATGA  912

Query  3224  TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  3283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  972

Query  3284  AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  3343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  1032

Query  3344  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  3403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1033  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  1092

Query  3404  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGA  3463
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  1093  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAANA  1152

Query  3464  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  3523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1153  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  1212

Query  3524  TATTTCTTAAAAAATATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  3583
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1213  TATTTCTTAAAAANTATTCTTAAATTGGATTCTAGAGGGCCCGTTTAAACCCGCTGATCA  1272

Query  3584  GCCTCGACTGT-GCCTTCTAGTTGCCAGCCATCTGTT-GTTTGCCCCTCCCCC-GTGCCT  3640
             |||||||||||| |||||||||||||| |||||||| | |||||||| |||||  | ||||
Sbjct  1273  GCCTCGACTGTNGCCTTCTAGTTGCCNGCCATCTGNTGGTTTGCCCNTCCCCNNGGGCCT  1332

Query  3641  TCCTTG-ACCCTGGAA-GGTGCCACTCCCACTGT-CCTTTCCTAATAAAA-TGAGG-AAA  3695
             |||||| ||||||||| ||||||    ||    ||  ||||||| || |||| ||||| |||
Sbjct  1333  TCCTTGGACCCTGGAAAGGTGCCCNNCCNNNTGGNCCTTTCCNAAAAAAANTGAGGNAAA  1392

Query  3696  TTGCAT-CGCATTG  3708
             |||||| | |||||
Sbjct  1393  TTGCATNCCCATTG  1406
```

Seq482

Score	Expect	Identities	Gaps	Strand
2427 bits(1314)	0.0	1354/1380(98%)	2/1380(0%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  13    GCGGNNACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  72

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  132

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  192

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||  ||| | 
Sbjct  193   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACAGCTCTTA  252

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  312

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  372

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  432

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  433   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  492

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  552

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  612

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  672

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  732

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGT  792

Query  3740  ggggtgggaCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   GGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  852

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  853   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  912

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  972

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  1032

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTT  4039
             ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Sbjct  1033  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTT  1092

Query  4040  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  4099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  1152

Query  4100  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  4159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1153  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  1212

Query  4160  GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA  4219
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Sbjct  1213  GGACTCTTGTTCCAAACTGGAACAACNNTCAACCCTATCTCGGTCTATTCTTTTGATTTA  1272

Query  4220  TAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTT  4279
             ||||||| ||||   ||||||||||||||||||||||  ||||||||||||||||||||
Sbjct  1273  TAAGGGAATTTGCCAATTTCGGCCTATTGGTTAAAAAATGNNCTGATTTAACAAAAATTT  1332

Query  4280  AACGCGAATTAATTCTGTGGAATGTGT-GTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCC  4338
             |  ||||||||||||  ||||||| || |||  |||||||| |||||||| ||||||  |||
Sbjct  1333  ACNGCGAATTAATTCCNTGGAATG-GTGGTCNNTTAGGGTNTGGAAAGTNCCCAGGNTCC  1391
```

6.2.7 L405A/I543A/V546A – Mut2

T7 promoter (F)

Score	Expect	Identities	Gaps	Strand
2518 bits(1363)	0.0	1379/1391(99%)	0/1391(0%)	Plus/Plus

```
Query  914   CTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGGTG  973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16    CTGGTACCGCCACCATGGCATGGAGCCACCCACAGTTCGAGAAGGGAGGTGGAAGCGGTG  75

Query  974   GAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGACA  1033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  76    GAGGCTCAGGAGGCAGCGCATGGTCCCACCCCCAGTTTGAAAAGCTTGCCACCATGGACA  135

Query  1034  AAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTCTG  1093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  136   AAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGGCAAGCTGGAACTGTCTG  195

Query  1094  GGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGACG  1153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  196   GGTGCGAACAGGGCCTGCACGAGATCAAGCTGCTGGGCAAAGGAACATCTGCCGCCGACG  255

Query  1154  CCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCA  1213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  256   CCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCA  315

Query  1214  CCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAG  1273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  316   CCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAG  375

Query  1274  CCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGC  1333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  376   CCCTGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAACTGC  435

Query  1334  TGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGGCA  1393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  436   TGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACCAGCAGCTGGCCGCCCTGGCCGGCA  495

Query  1394  ATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCTGA  1453
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  496   ATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATTCTGA  555

Query  1454  TCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCTCG  1513
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  556   TCCCCTGCCACCGGGTGGTGTCTAGCTCTGGCGCCGTGGGGGGCTACGAGGGCGGGCTCG  615

Query  1514  CCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGCG  1573
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  616   CCGTGAAAGAGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGCG  675

Query  1574  AATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAATG  1633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  676   AATTCATGTCTTCCAGTAATGTCGAAGTTTTTATCCCAGTGTCACAAGGAAACACCAATG  735

Query  1634  GCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAGTT  1693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  736   GCTTCCCCGCGACAGCTTCCAATGACCTGAAGGCATTTACTGAAGGAGCTGTGTTAAGTT  795

Query  1694  TTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACCAG  1753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  796   TTCATAACATCTGCTATCGAGTAAAACTGAAGAGTGGCTTTCTACCTTGTCGAAAACCAG  855

Query  1754  TTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCATCC  1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  856   TTGAGAAAGAAATATTATCGAATATCAATGGGATCATGAAACCTGGTCTCAACGCCATCC  915

Query  1814  TGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGATC  1873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  916   TGGGACCCACAGGTGGAGGCAAATCTTCGTTATTAGATGTCTTAGCTGCAAGGAAAGATC  975

Query  1874  CAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAAT  1933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  976   CAAGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAAT  1035

Query  1934  GTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAA  1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1036  GTAATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAA  1095

Query  1994  ACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaCG  2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1096  ACTTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAACG  1155

Query  2054  AACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTG  2113
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  1156  AACGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCANACTCCAAGGTTG  1215

Query  2114  GAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGG  2173
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Sbjct  1216  GAACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAAAAAAAGGACTAGTATAGGAATGG  1275

Query  2174  AGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAA  2233
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Sbjct  1276  AGCTTATCACTGATCCTTCCATCTTGTTCTTGGATGAACCTACAACTGGCTTAGACTCAA  1335

Query  2234  GCACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCA  2293
             |  |||||||||||||||||||||||||||||||||||||| | ||||| | ||||||
Sbjct  1336  GNCCAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAANCNGGGACAANCAATCA  1395

Query  2294  TCTTCTCCATT  2304
               |||  |||||
Sbjct  1396  NCTTNNCCATT  1406
```

SeqF1

Score	Expect	Identities	Gaps	Strand
2409 bits(1304)	0.0	1342/1364(98%)	6/1364(0%)	Plus/Plus

```
Query  1876  AGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGT  1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  21    AGTGGATTATCTGGAGATGTTCTGATAAATGGAGCACCACGACCTGCCAACTTCAAATGT  80

Query  1936  AATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAAC  1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  81    AATTCAGGTTACGTGGTACAAGATGATGTTGTGATGGGCACTCTGACGGTGAGAGAAAAC  140

Query  1996  TTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGaaaaaaaCGAA  2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  141   TTACAGTTCTCAGCAGCTCTTCGGCTTGCAACAACTATGACGAATCATGAAAAAAACGAA  200

Query  2056  CGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGA  2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  201   CGGATTAACAGGGTCATTCAAGAGTTAGGTCTGGATAAAGTGGCAGACTCCAAGGTTGGA  260

Query  2116  ACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAG  2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  261   ACTCAGTTTATCCGTGGTGTGTCTGGAGGAGAAAGAAAAAGGACTAGTATAGGAATGGAG  320

Query  2176  CTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGC  2235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  321   CTTATCACTGATCCTTCCATCTTGTTCTTGGATGAGCCTACAACTGGCTTAGACTCAAGC  380

Query  2236  ACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATC  2295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  381   ACAGCAAATGCTGTCCTTTTGCTCCTGAAAAGGATGTCTAAGCAGGGACGAACAATCATC  440

Query  2296  TTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTG  2355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  441   TTCTCCATTCATCAGCCTCGATATTCCATCTTCAAGTTGTTTGATAGCCTCACCTTATTG  500

Query  2356  GCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCA  2415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  501   GCCTCAGGAAGACTTATGTTCCACGGGCCTGCTCAGGAGGCCTTGGGATACTTTGAATCA  560

Query  2416  GCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAAT  2475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  561   GCTGGTTATCACTGTGAGGCCTATAATAACCCTGCAGACTTCTTCTTGGACATCATTAAT  620

Query  2476  GGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATA  2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  621   GGAGATTCCACTGCTGTGGCATTAAACAGAGAAGAAGACTTTAAAGCCACAGAGATCATA  680

Query  2536  GAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCC  2595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  681   GAGCCTTCCAAGCAGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCC  740

Query  2596  TCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAG  2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  741   TCCTTCTACAAAGAGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAAGAAG  800

Query  2656  AAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGG  2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  801   AAGATCACAGTCTTCAAGGAGATCAGCTACACCACCTCCTTCTGTCATCAACTCAGATGG  860

Query  2716  GTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAGATC  2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  861   GTTTCTAAGCGTTCATTCAAAAACTTGCTGGGTAATCCCCAGGCCTCTATAGCTCAGATC  920

Query  2776  ATTGTCACAGTCGTACTGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGAT  2835
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  921   ATTGTCACAGTCGTAGCGGGACTGGTTATAGGTGCCATTTACTTTGGGCTAAAAAATGAT  980

Query  2836  TCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGTTTC  2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  981   TCTACTGGAATCCAGAACAGAGCTGGGGTTCTCTTCTTCCTGACGACCAACCAGTGTTTC  1040

Query  2896  AGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATAC  2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1041  AGCAGTGTTTCAGCCGTGGAACTCTTTGTGGTAGAGAAGAAGCTCTTCATACATGAATAC  1100

Query  2956  ATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTA  3015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1101  ATCAGCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTA  1160

Query  3016  CCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTG  3075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1161  CCCATGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTG  1220

Query  3076  AAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGAT-GGTGGCTTATTC  3134
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  1221  AAGCCAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGGTGGCTTATTC  1280

Query  3135  AGCCAGTTCCATGG-CACTGG-CCATAGCAGCAGGTCAGAGTGT-GGTTTCTGTAGCAAC  3191
              |||||||||||||| |||||| ||||||||||||||||||     |||||| | |||||
Sbjct  1281  AGCCAGTTCCATGGGCACTGGNCCATAGCAGCAGGTCAGANNNNGGGTTTCNGNAGCAAN  1340

Query  3192  AC-TTCTCATGACCATCTGTTTTGTGTTTATG-ATGATTTTTTC  3233
               | ||||||||||  ||||||||  ||||||  |  ||||||||
Sbjct  1341  NCNTTCTCATGACCGCCTGTTTTGCCTTTATGNANAATTTTTTC  1384
```

SeqF2

Score	Expect	Identities	Gaps	Strand
2386 bits(1292)	0.0	1338/1365(98%)	2/1365(0%)	Plus/Plus

```
Query  2324  TCTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  2383
             |||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14    TCTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGC  72

Query  2384  CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  2443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  73    CTGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATA  132

Query  2444  ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  2503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  133   ACCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACA  192

Query  2504  GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  2563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193   GAGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCA  252

Query  2564  TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  2623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  253   TAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAAT  312

Query  2624  TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  2683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  313   TACATCAACTTTCCGGGGGTGAGAAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCT  372

Query  2684  ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  2743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  373   ACACCACCTCCTTCTGTCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAACTTGC  432

Query  2744  TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTA  2803
              ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Sbjct  433   TGGGTAATCCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTAGCGGGACTGGTTA  492

Query  2804  TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  2863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  493   TAGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGG  552

Query  2864  TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  2923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  553   TTCTCTTCTTCCTGACGACCAACCAGTGTTTCAGCAGTGTTTCAGCCGTGGAACTCTTTG  612

Query  2924  TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  2983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  613   TGGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGTCATCTT  672

Query  2984  ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  3043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  673   ATTTCCTTGGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATAT  732

Query  3044  TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  3103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  733   TTACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTA  792

Query  3104  TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  3163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  793   TGATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCCATGGCACTGGCCATAGCAG  852

Query  3164  CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTTGTGTTTATGA  3223
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||  |||||||
Sbjct  853   CAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCGCCTGTTTTGCCTTTATGA  912

Query  3224  TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  3283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  913   TGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC  972

Query  3284  AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  3343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  973   AGTACTTCAGCATTCCACGATATGGATTTACGGCTTTGCAGCATAATGAATTTTTGGGAC  1032

Query  3344  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  3403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1033  AAAACTTCTGCCCAGGACTCAATGCAACAGGAAACAATCCTTGTAACTATGCAACATGTA  1092

Query  3404  CTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGA  3463
              ||||||||.||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1093  CTGGCGAANAATATTTGGTAAAGCAGGGCATCGATCTCTCACCCTGGGGCTTGTGGAAGA  1152

Query  3464  ATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  3523
              ||||||||||||||||||.|||| ||||||||||||||||||||||||||||||||||||
Sbjct  1153  ATCACGTGGCCTTGGCTTGGATGAATGTTATTTTCCTCACAATTGCCTACCTGAAATTGT  1212

Query  3524  TATTTCTTAAAAAATATTCTTAAATTGG-ATTCTAGAGGGCCCGTTTAAACCCGCTGATC  3582
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  1213  TATTTCTTAAAAAATATTCTTAAATTGGAATTCTAGAGGGCCCGTTTAAACCCGCTGATC  1272

Query  3583  AGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTC  3642
              |||||||||.|.|||||   |||||| |||||| ||||||||||||||||||||||||||
Sbjct  1273  AGCCTCGACGGGGCCTTNNNGTTGCCNGCCATCGGTTGTTTGCCCCTCCCCCGTGCCTTC  1332

Query  3643  CTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA  3687
              |||||||||||||||.|||   ||| ||  |||||||.|| ||||
Sbjct  1333  CTTGACCCTGGAAGGGGCCCNNCCCCCTNNCCTTTCCCAAAAAAA  1377
```

Seq482

Score	Expect	Identities	Gaps	Strand
2398 bits(1298)	0.0	1352/1383(98%)	8/1383(0%)	Plus/Plus

```
Query  2960  GCGGATACTACAGAGTGTCATCTTATTTCCTTGGAAAACTGTTATCTGATTTATTACCCA  3019
             ||||  ||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  12    GCGGNNACTACAGAGTGTCATCTTATTTCCTTGGNAAACTGTTATCTGATTTATTACCCA  71

Query  3020  TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  72    TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC  131

Query  3080  CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132   CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA  191

Query  3140  GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  3199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  192   GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA  251

Query  3200  TGACCATCTGTTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  3259
             |||||  ||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  252   TGACCGCCTGTTTTGCCTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA  311

Query  3260  TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  312   TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT  371

Query  3320  TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  372   TGCAGCATAATGAATTTTTGGGACAAAACTTCTGCCCAGGACTCAATGCAACAGGAAACA  431

Query  3380  ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  432   ATCCTTGTAACTATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC  491

Query  3440  TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  492   TCTCACCCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTTCC  551

Query  3500  TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  552   TCACAATTGCCTACCTGAAATTGTTATTTCTTAAAAAATATTCTTAAATTGGATTCTAGA  611

Query  3560  GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  612   GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT  671

Query  3620  GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672   GTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC  731

Query  3680  TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTgggggggt  3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  732   TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGGT  791

Query  3740  ggggtgggggCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  792   GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  851

Query  3800  GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  3859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  852   GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC  911

Query  3860  CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  3919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  912   CACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC  971

Query  3920  GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  3979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  972   GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC  1031

Query  3980  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTT  4039
             ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Sbjct  1032  ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTT  1091

Query  4040  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG  4099
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  1092  AGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAATGGG  1151

Query  4100  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT  4159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1152  CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGN  1211

Query  4160  GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA  4219
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1212  GAACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA  1271

Query  4220  TAAGGGATTTTGGGGA-TTTC-GGCCTATTGGTT-AAAAAATGAGCTGATTTAACAAAAA  4276
             ||||||||||||   | |||| ||||||||||||| ||||||||||||||||||||||||
Sbjct  1272  TAAGGGATTTTGCCNAATTTCNGGCCTATTGGTTAAAAAAATGAGCTGATTTAACAAAAA  1331

Query  4277  TTTAACG-CGAATTAATTCTGTGGAATGTGT-GTCAG-TTAGGGTGTGGAAA-GTCCCC-  4331
             ||||||| ||||||||||   |||||  |  | |||||  |||||  |||||||  ||||||
Sbjct  1332  TTTAACGCCGAATTAATTNNNTGGAANGGNTNGTCAGNTTAGGNNGTGGAAAAGTCCCCC  1391

Query  4332  AGG  4334
             |||
Sbjct  1392  AGG  1394
```